THE ENHANCED U.S. NCAP: FIVE YEARS LATER

Brian T. Park
Taryn E. Rockwell
National Highway Traffic Safety Administration
United States of America

Lauren A. Collins
Christina S. Smith
Mai Lan Aram
Alpha Technology Associate, Inc.
United States of America

Paper Number 15-0314

ABSTRACT

The U.S. New Car Assessment Program (NCAP) encourages vehicle manufacturers to make safety improvements to their vehicles through its award-winning consumer information-based 5-Star Safety Ratings Program. Occupant injury readings have decreased and star ratings have generally improved since the program was enhanced in the 2011 Model Year (MY). This paper summarizes vehicle crash test results for the five MYs since the program was last upgraded to demonstrate how quickly vehicles have been redesigned to achieve high ratings under the more stringent requirements. As a result, most vehicles are achieving 4- or 5-star ratings. Though there are still vehicles the agency has tested that do not achieve the highest ratings, the performance of the majority of vehicles tested under the enhanced program is excellent. This sets the stage for the agency to begin exploring the possibility of making additional changes to the current program to spur even further vehicle safety improvements through market forces and consumer demand. The NCAP's crash test data (specifically, occupant injury data) and star ratings derived from those data will be used throughout this study. Occupant performance from the first year of the enhanced program will be compared to more recent results. A comparative analysis of paired data for vehicles that have been tested and retested under the enhanced NCAP will also be shown. These analyses will serve to demonstrate the effectiveness of the program in encouraging vehicle manufacturers to make immediate design changes that improve the occupant protection afforded by their vehicles.

LIMITATIONS OF THIS STUDY

The crash test data included in this study spans MYs 2011 to 2015; however, MY 2015 data is limited as testing was not completed in time for all results to be included in this publication. Also, the study is limited by the vehicles the NCAP selected each MY for testing. Though these tests encompass results from over 85% of the vehicles (by projected sales volume) in the U.S. fleet each MY they do not represent the composition of the entire fleet (NHTSA 2013d). Therefore, the authors do not make any further projections about overall fleet safety. Furthermore, this study also does not include an analysis of how the probabilities of injury risk measured in NCAP tests relate to the reduction of real-world injuries and fatalities.

HISTORY OF THE U.S. NCAP

The National Highway Traffic Safety Administration (NHTSA) established the NCAP in 1978 in response to Title II of the Motor Vehicle Information and Cost Savings Act of 1972. At the time, the agency only assessed vehicles for their occupant protection in frontal impacts. Though stars were not used at the time, it has been estimated that less than 30% of those tested vehicles would have been able to achieve the highest possible ratings (4 or 5 stars) for the driver (2008a). After several years and the addition of a side moving deformable barrier (MDB) impact crash test for MY 1997 vehicles, and a rollover resistance test in MY 2001, the NCAP began soliciting comments for a program overhaul in 2007 (2007a). The agency began testing and rating vehicles under an enhanced program (NHTSA 2008a) in 2010 beginning with MY 2011 vehicles. Under the enhanced program, the agency continues to conduct a 35 mph (56 km/h) vehicle frontal impact rigid barrier test and a 38.5 mph (62 km/h) MDB side impact vehicle test, but now uses newer side-impact test dummies compared to the previous program. The enhanced program now also includes a 20 mph (32 km/h) vehicle side impact pole test. For each test type, injury data from each occupant is

collected from different body regions and is incorporated into a single combined probability of injury. An occupant's combined probability of injury is divided by a baseline risk of injury and the resulting relative risk score, or RRS, is converted to a star rating for that occupant. The results of all three crash tests are then merged with the results from a rollover resistance test, which remained unchanged from the previous version of the program, in order to provide consumers with a single, combined vehicle rating known as the Overall Vehicle Score (OVS) (NHTSA 2008a).

The agency examined the performance of vehicles tested soon after the enhanced program was initially launched and was encouraged that those results showed lower levels of injury when compared to the performance of vehicles tested under the previous version of the program (Park 2011). After five MYs, the agency believes that sufficient data has been collected to adequately assess occupant performance in modern vehicles under the NCAP. This paper will present the results of this analysis.

Since the rollover resistance test remained the same under the enhanced program and this paper focuses only on crashworthiness results, the authors will not explore the results of that testing mode over the past five MYs. Likewise, since rollover results are needed to calculate the OVS, this paper will not discuss how that particular metric has changed. Each section below will briefly describe the test setup and protocol for each crash test mode before describing the results of the study. If desired, additional procedural details can be obtained from the individual test protocols (NHTSA 2013a, b, and c), and more detailed information pertaining to the injury risk curves and ratings system can be found in the appendices of the "Final decision" notice (NHTSA 2008a).

METHODOLOGY OF ANALYSES

The purpose of this study was to determine the effectiveness of the program over the five MYs since the enhancements were implemented, primarily by comparing crash test data from previous versions of make models selected and tested in the program with those of a later version. To accomplish this, NCAP test data was analyzed in two ways.

First, for each test mode, aggregate crash test results from the beginning of the enhanced NCAP were compared to the most recent data collected. Unpaired t-tests were conducted to determine whether occupant injury probability and star ratings from MY 2011vehicles were different from those collected for MY 2014-2015 vehicles. The last two years of test data were combined to ensure sufficient data points for analysis since the full set of 2015 data was not available at the time of publication. For ease of discussion, this analysis will be referred to as the unpaired analysis. Both star ratings and the occupant combined injury probabilities were analyzed because it was thought that the two could show different nuances in the data.

Second, vehicles that were tested more than once from the 2011 to 2015 MYs were identified and isolated to perform a similar analysis. At the beginning of each MY, the NCAP solicits information about light vehicle production from each vehicle manufacturer. In these requests, manufacturers are given the opportunity to inform the agency about the vehicles they plan to produce in the next MY. Some vehicle designs are considered identical from one MY to the next. When this is the case, the NCAP is able to "carry over" a given set of vehicle ratings to the next MY. In other cases, design changes are made that only affect the performance of a vehicle in one test mode. For example, changes may be made to the design of a frontal air bag which could result in retesting the vehicle for the frontal rigid barrier test, but not for the side impact test modes.

The authors examined data from the "first" and "last" times a make and model was tested under the enhanced NCAP and, using paired t-tests, tried to determine whether design changes between those two versions resulted in measurable improvements in the crash protection afforded to the vehicle occupants. Some vehicles may have been tested more than twice over this time period, such that other test data fell in the interim – either after the "first" test or before the "last" test. In these cases, manufacturers may have incorporated one or more interstitial design changes in vehicles that were not necessarily successful in achieving better star ratings, which resulted in retesting those vehicles several times during the course of the enhanced program. For ease of calculation and discussion, the authors decided to exclude this interim data and focus only on the first and last tests from the past five MYs. By selecting the latest version tested, it was expected that manufacturers' best efforts to date would be captured. This second

analysis will be referred to as the paired analysis. Test results from the first time a given vehicle model was tested under the enhanced program were compared to those from the most recent time that vehicle was retested.

In addition, data was analyzed based on changes to the ratings assigned, which may or may not be represented by the same vehicle make and model from year to year. This is due to several factors. For instance, vehicle manufacturers may change and/or rebrand particular vehicle names from one year to the next. A vehicle model may also be referred to as something new from a marketing perspective, but in actuality, it is identical to a model that was previously produced. In addition, vehicles that a manufacturer claims are corporate twins in one MY may not be designated as such in subsequent years.

Lastly, some vehicles demonstrated an increase in risk of occupant injuries (or a subsequent decrease in star ratings) after a redesign. No additional effort was made to determine how or why these vehicles performed more poorly after a redesign.

RESULTS

Frontal NCAP – Rigid Barrier Test

The NCAP's frontal star ratings are based on the performance of two Hybrid III dummies installed in the first row of a vehicle. The driver is a 50th percentile male and the right front passenger is a 5th percentile female. For this test, the vehicle impacts a rigid wall at 35 mph (56 km/h) (NHTSA 2013a). Readings from the dummies' heads, necks, chests, and legs are used to calculate combined probabilities of AIS3+ injury, which are then divided by a baseline risk of injury and converted to star ratings (NHTSA 2008a). For ease of discussion, this paper will primarily focus on the probabilities of injuries recorded by the test dummies and the star ratings assigned to them, and will not explain the calculations involved, as these are presented in detail in the 2008 "Final decision" notice (2008a).

A summary of the average probabilities of injury and star ratings over the years for the driver and front passenger dummies in the frontal NCAP test may also be found in the Appendix in Tables A1-A4.

Comparison of unpaired vehicle data – 2011 versus 2014/2015 (unpaired analysis) The average combined injury probabilities and star ratings for the driver dummy did not change much from MY 2011 to MY 2014/2015. There was no statistical difference found at the 95% confidence interval as evidenced by the P-value between the MY 2011 and MY 2014/2015 data sets when considering the occupants' combined injury probabilities and star ratings. However, differences between MY 2011 and MY 2014/2015 average combined injury probabilities and star ratings for the right front passenger were found to be statistically significant at the 95% confidence interval. The combined injury probabilities showed a significant decrease, and consequently, the star rating significantly increased. Table 1 details the results of this analysis.

Table 1.
Results of Frontal NCAP Unpaired Analysis

	Frontal NCAP Driver				**Frontal NCAP Front Passenger**			
	Comb. Injury Prob.		***Star Rating***		***Comb. Injury Prob.***		***Star Rating***	
Data Set	**2011**	**2014/15**	**2011**	**2014/15**	**2011**	**2014/15**	**First Test**	**Last Test**
t-stat	0.928		0.386		4.219		4.532	
Two tailed P-value	0.355		0.700		< 0.0001		< 0.0001	
Mean	0.117	0.111	4.28	4.33	0.151	0.117	3.55	4.24
Standard Deviation	0.036	0.036	0.81	0.73	0.051	0.042	0.92	0.82
N	64	66	64	66	64	66	64	66

Comparison of paired vehicle data – 2011 versus most recent test (paired analysis) While driver performance in frontal NCAP has not improved significantly on average from MY 2011, the paired analysis shows that when manufacturers make changes to specific vehicles, significant improvements are seen in both combined injury probabilities and star ratings. These differences were significant at the 95% confidence level. Likewise, differences in the most updated vehicles' right front passenger combined injury probabilities and star ratings were

also statistically significant at the 95% confidence level. Both the driver and right front passenger dummies' combined injury probabilities showed a significant decrease, and consequently, the star ratings significantly increased. Table 2 illustrates the results for this analysis.

Table 2.
Results of Frontal NCAP Paired Analysis

	Frontal NCAP Driver				**Frontal NCAP Front Passenger**			
	Comb. Injury Prob.		***Star Rating***		***Comb. Injury Prob.***		***Star Rating***	
Data Set	**First Test**	**Last Test**	**First Test**	**Last Test**	**First Test**	**Last Test**	**First Test**	**Last Test**
t-stat	4.443		3.244		7.889		7.496	
Two tailed P-value	< 0.0001		0.0018		< 0.0001		< 0.0001	
Mean	0.121	0.103	4.19	4.49	0.158	0.118	3.41	4.18
Standard Deviation	0.040	0.023	0.84	0.58	0.054	0.038	0.93	0.71
N	73		73		73		73	

Side NCAP – MDB Test

The side MDB NCAP test is conducted by towing an MDB into the driver's side of a stationary vehicle crabbed at a 27 degree angle at 38.5 mph (62 km/h) (NHTSA 2013b). An ES-2re 50th percentile male dummy is seated in the driver's seat, and a 5th percentile female SID-IIs dummy is seated in the rear seat behind the driver. Readings from the driver dummies' heads, chests, abdomens, and pelvises, and the rear passenger dummies' heads and pelvises, are used to calculate combined probabilities of AIS3+ injury, which are then divided by a baseline risk of injury and converted to star ratings (NHTSA 2008a).

A summary of the average injury probabilities over the years for the driver and rear passenger dummies in the side MDB test may be found in the Appendix in Tables A5-A8.

Comparison of unpaired vehicle data – 2011 versus 2014/2015 (unpaired analysis) As shown in Table 3, the differences between both the star ratings and the combined injury probabilities in the 2011 data and the 2014/2015 data were found to be statistically significant at the 95% confidence interval for both dummies in the side MDB test. Both the driver and rear passenger dummies' combined injury probabilities showed a significant decrease, and consequently, the star ratings significantly increased. Five-star ratings dominate the 2014/2015 data set; 89% of driver ratings and 96% of rear passenger ratings were five stars. These figures can be compared to 67% of driver ratings and 68% of rear passenger ratings that achieved five stars in the 2011 data set.

Table 3.
Results of Side MDB NCAP Unpaired Analysis

	Side MDB NCAP Driver				**Side MDB NCAP Rear Passenger**			
	Comb. Injury Prob.		***Star Rating***		***Comb. Injury Prob.***		***Star Rating***	
Data Set	**2011**	**2014/15**	**2011**	**2014/15**	**2011**	**2014/15**	**2011**	**2014/15**
t-stat	3.820		3.518		3.353		3.585	
Two tailed P-value	0.0002		0.0006		0.0010		0.0005	
Mean	0.099	0.057	4.44	4.87	0.087	0.047	4.37	4.88
Standard Deviation	0.088	0.034	0.96	0.41	0.089	0.051	1.04	0.59
N	64	75	64	75	62	75	62	75

Comparison of paired vehicle data – 2011 versus most recent test (paired analysis) For the paired analysis, Table 4 shows the differences between the combined injury probabilities and star ratings in the two sets of data ("first" test versus "last" test) were found to be significant at the 95% confidence interval. Both the driver and rear

passenger dummies' combined injury probabilities showed a significant decrease, and consequently, the star ratings significantly increased.

Table 4.
Results of Side MDB NCAP Paired Analysis

	Side MDB NCAP Driver				**Side MDB NCAP Rear Passenger**			
	Comb. Injury Prob.		***Star Rating***		***Comb. Injury Prob.***		***Star Rating***	
Data Set	**First Test**	**Last Test**	**First Test**	**Last Test**	**First Test**	**Last Test**	**First Test**	**Last Test**
t-stat	3.884		4.323		4.073		3.915	
Two tailed P-value	0.0002		< 0.0001		0.0001		0.0002	
Mean	0.093	0.062	4.38	4.79	0.093	0.046	4.32	4.86
Standard Deviation	0.083	0.052	0.91	0.58	0.095	0.048	1.10	0.51
N	72		72		72		72	

Side NCAP – Pole Test

The side impact pole test, which was new to the NCAP beginning with the enhanced program, is conducted by towing a test vehicle angled at 75 degrees sideways at 20 mph (32 km/h) into an 8 in. (20 cm) diameter rigid pole (NHTSA 2013c). To assess the maximum injury potential in this type of crash, the test protocol requires that the pole be aligned with the center of gravity (CG) of the SID-IIs 5th percentile female driver dummy's head. Readings from the dummies' heads and pelvises are used to calculate combined probabilities of AIS3+ injury, which are then divided by a baseline risk of injury and converted to star ratings (NHTSA 2008a).

A summary of the average injury probabilities and star ratings over the years for the driver dummy in the side pole test may be found in the Appendix as Tables A9 and A10.

Comparison of unpaired vehicle data – 2011 versus 2014/2015 (unpaired analysis) The driver dummy in the side NCAP pole test has seen a large increase in average star ratings since the beginning of the enhanced program in 2011. Both the dummy and the crash test type were completely new beginning with that MY. Despite the obvious increase in pole test driver star ratings over the past few years, a t-test to determine significance was performed comparing the first MY of data to the most recent. The results of this analysis proved statistical significance for the decrease in the combined probability of injury and for the increase in the star rating at the 95% confidence level.

Table 5.
Results of Side Pole NCAP Unpaired Analysis

	Side Pole NCAP Driver			
	Comb. Injury Prob.		***Star Rating***	
Data Set	**2011**	**2014/15**	**2011**	**2014/15**
t-stat	4.229		4.448	
Two tailed P-value	< 0.0001		< 0.0001	
Mean	0.137	0.063	4.10	4.84
Standard Deviation	0.145	0.034	1.35	0.44
N	62	74	62	74

Comparison of paired vehicle data – 2011 versus most recent test (paired analysis) For the paired analysis, the data shows that manufacturers' efforts to improve performance were extremely successful. The t-test analysis shows that between the "first" and "last" vehicle tests, the reduction in probability of injury and the increase in star ratings are both significant at the 95% confidence level. The driver dummies' combined injury probabilities showed a significant decrease, and consequently, the star rating significantly increased.

Table 6.
Results of Side Pole NCAP Paired Analysis

	Side Pole NCAP Driver			
	Comb. Injury Prob.		*Star Rating*	
Data Set	**First Test**	**Last Test**	**First Test**	**Last Test**
t-stat	4.518		5.688	
Two tailed P-value	< 0.0001		< 0.0001	
Mean	0.181	0.065	3.79	4.81
Standard Deviation	0.216	0.033	1.46	0.43
N	70		70	

DISCUSSION

Percentages listed in the below analyses are derived from the data found in the Appendix.

Unpaired Analysis

It is expected that the driver data in the frontal NCAP test would not show significance when comparing the MY 2011 data to the most recent years' data because the dummy type and seating method remained the same in the enhanced program even though the method used to assess the driver dummy's injuries was changed. The 50th percentile Hybrid III dummy has been a fixture as the driver in the frontal NCAP test since 1986 (1995). These circumstances may have given manufacturers a head start in developing design strategies for vehicles tested under the requirements of the enhanced program.

When the enhanced program began (MY 2011), the right front occupant in the frontal crash test had a larger proportion of two- and three-star ratings than the frontal driver did; nearly 38% of right front passenger ratings were less than four stars as opposed to 16% of driver ratings. By MY 2014/2015, these percentages had decreased to 12% for the front passenger and 9% for the driver. The occupants that showed the greatest improvement since the enhanced program began in terms of average star ratings were the Hybrid III 5th percentile female in the right front passenger seat of the frontal test and the 5th percentile SID-IIs driver in the pole test. Since the enhanced program's inception, these occupants experienced a 19% and 18% increase in average star ratings, respectively. The use of 5th percentile female dummies was new to the NCAP in MY 2011. The pole test configuration itself was also completely new. Manufacturers have developed countermeasures to reduce injury risk in these crashes. For instance, the combined probability of injury for the 5th percentile female in the pole test has been reduced by more than half (54%) since MY 2011. For the side MDB test, the probability of injury in the 5th percentile female SID-IIs, seated in the rear seat, reduced by half – more or less (47%), and star ratings increased by 12% on average.

The 50th percentile male ES-2re dummy was also new for the enhanced program. Similar to the results of the SID-IIs in the side MDB test, the ES-2re dummy showed a marked decrease in the combined probability of injury (42%) with only a moderate increase in average star ratings (10%).

The NCAP also tests vehicles beyond the requirements of the Federal Motor Vehicle Safety Standards (FMVSSs), and this is accomplished in several ways. For instance, the NCAP tests vehicles that are beyond the weight requirements set forth in various FMVSSs. This allows the program to select and test vehicles with up to a 10,000 lb (4,536 kg) Gross Vehicle Weight Rating (GVWR) to ensure that some popular consumer vehicles are subject to full-scale crash testing when they may not have been otherwise. Hence, the NCAP test results of certain vehicles may highlight the need for further refinements perhaps in the restraint systems to achieve better NCAP performance. For instance, side air bags, particularly chest and pelvis air bags, are now seen more often in larger vehicles. Note that risk of head injury in the side MDB test is very low and has been, at least since the enhanced program's inception (Table A5). Prior to the enhanced program, side curtain air bags were not necessarily needed to achieve top ratings for these vehicles, but the addition of the side pole test has challenged manufacturers to include safety equipment for head and chest protection on vehicles, including those with GVWRs that exceed FMVSSs. In the MYs during which the enhanced program was formally announced (MY 2009), no vehicles with GVWR between 8,500 lb and 10,000 lb (3,856 kg to 4,536 kg) offered side air bags to protect the occupant's chest or pelvis, and only 38% of these

models offered head protection. Now, many more vehicle models in the MY 2015 heavy vehicle fleet offer side air bags, and all models offer head protection, with 47% of them having head air bags as standard equipment. Additionally, 73% of models offer chest and/or pelvis protection as either optional or standard equipment.

Paired Analysis

Injury probabilities and star ratings examined in the paired analysis showed promising results for all occupants in the NCAP's crash tests.

While the average combined injury probability for the frontal crash test driver decreased by only 5% in the unpaired analysis, it decreased by 15% when looking at the paired vehicle tests; the decrease was sufficient to achieve statistical significance. Accordingly, the average star ratings for the frontal crash test driver increased by 1% in the unpaired analysis versus an increase of 7% in the paired analysis. The frontal crash test passenger's combined injury probability decreased by 25% and the star ratings increased by 23% in the paired study. This provides evidence that manufacturers are responding to frontal NCAP crash test results and further support the hypothesis that manufacturers are making specific design changes to improve occupant performance in vehicles with lower NCAP star ratings even when the larger population of frontal NCAP vehicles shows generally good performance.

The side pole test driver in this paired analysis experienced the largest decrease (64%) in average combined probability of injury of any occupant in either analysis and a 27% increase in average star ratings, an average of greater than one whole star, which was the largest of any occupant's average star rating increase. In fact, a large number of pole test ratings increased by three (n=10, 14%) and even four (n=6, 9%) stars after redesign. In the same analysis, the driver in the side MDB test experienced a 33% reduction in average combined injury probability. The reductions in both MDB and pole test driver combined injury probabilities demonstrate that the driver's position in side impact tests is becoming safer for both the 50th percentile male and the 5th percentile female. Also, it appears countermeasures can be put in place to improve injury readings for both the 50th percentile male and the 5th percentile female. The rear passenger in the MDB test paired data set saw a decrease in average combined injury probability of 51% and star ratings increased 13%. No ratings increased by four stars, but unlike in the pole test, none of the rear passenger ratings were one-star initially, so a four-star increase would be unattainable. Of the ten vehicles in which the rear passenger received a two-star rating, eight vehicles were modified in the later version and achieved five-star ratings.

Though the specific details of a given vehicle's design changes are considered confidential information, there are some general trends in restraint and structural changes the authors can note that have occurred during the past five years of the enhanced program. To improve their vehicles' frontal crash performance, manufacturers made air bag design and deployment refinements, and modifications to advanced seat belt features such as load limiters and pretensioners. Manufacturers have also made front-end changes to better manage crash energy. Regarding the side impact crash mode, the presence of the 5th percentile female SID-IIs in the rear seat of each MDB test seems to have encouraged manufacturers to equip more rear seating positions with side torso air bags or larger torso-abdomen-pelvis (TAP) air bags. This equipment was not often seen prior to the enhanced program. Manufacturers have also opted to install TAP air bags in many front row seats to maximize vehicle performance in both side impact tests. Structural improvements to the side sills, pillars, and door trims have also been made to reduce intrusion into the occupant compartment and/or provide a load path through the vehicle, particularly for the pole test. Chamber locations in curtain air bags have been modified to protect both the 50th percentile male and the 5th percentile female in side crash scenarios.

Four of the vehicles included in the paired analysis had GVWRs between 8,500 lb and 10,000 lb (3,856 kg to 4,536 kg). In all four cases, the combined probabilities of injury substantially improved for the pole test between the first and last test. Three driver occupants experienced high head injury probabilities during the first test. In these cases, injury probabilities appeared to be later improved, in part by the addition of side curtain air bags. The fourth case involved a driver whose pelvic injury probability was high. Side torso air bags were added to this vehicle, among other improvements, and its star rating increased dramatically.

Occupant protection performance in vehicle crash tests is complex. There is no single way to reach a five-star rating, and vehicle manufacturers employ a variety of techniques to achieve the top ratings. In addition, manufacturers must take into account other factors such as vehicle weight, style, and comfort. However, because of improvements made to vehicles, most manufacturers were successful in their efforts to decrease probability of injury and therefore

increase star ratings. The paired analysis suggests that, in the majority of cases, when manufacturers choose to implement design changes, occupant safety is enhanced.

CONCLUSIONS

Two methods of analysis were used to compare changes to occupant injury risk under the enhanced NCAP. An unpaired analysis was undertaken to generally assess the state of injury risk and vehicle ratings in both the first MY of the enhanced program and the most recent MYs, 2014 and 2015. A paired analysis was then run to investigate whether vehicle redesigns were successful in reducing injury risk. All differences were statistically evaluated using t-tests, and for these analyses, significance is defined using a 95% confidence interval.

Driver data in the frontal NCAP test does not show a significant difference overall when comparing MY 2011 to 2014/2015. Likely, manufacturers had already had a head start in designing strategies to protect this occupant in this test mode, as the dummy, seating procedure, and test itself did not change under the enhanced program. However, when considering the paired analysis, driver results from the first and last tests of a vehicle, when one was retested, showed that injury risk was significantly decreased and star ratings significantly increased.

Passenger data in the frontal NCAP test, as well as driver and passenger data from the side impact tests (MDB and pole), show significant decreases in injury risk and, subsequently, significant increases in star rating. This is true for both the unpaired and paired analyses and suggests that manufacturers are successfully responding to the enhanced NCAP crash tests.

Vehicles with GVWRs between 8,500 lb to 10,000 lb (3,856 kg to 4,536 kg) are, in general, being fitted with additional safety equipment in recent MYs. The additions protect body regions which are evaluated in the enhanced NCAP's crash tests, particularly those in the side pole test. When side impact safety equipment was added to the four heavier vehicles included in the paired study, side pole star ratings increased substantially.

REFERENCES

"Consumer Information; New Car Assessment Program; Final decision notice," 73 Federal Register 134 (11 July 2008a), pp 40016- 40050.

"Consumer Information; New Car Assessment Program (NCAP); Notice of postponement of the implementation of enhancements to the New Car Assessment Program (NCAP)," 73 Federal Register 248 (24 December 2008b), pp 79206- 79207.

"Highway Safety: Reliability and Validity of DOT Crash Tests." GAO/PEMD-95-5. May 1995.

"Laboratory Test Procedure for New Car Assessment Program Frontal Impact Testing," Docket NHTSA-2013-0052, 18 September 2013a.

"Laboratory Test Procedure for New Car Assessment Program Side Impact Moving Deformable Barrier Testing." Docket NHTSA-2013-0052. 19 September 2013b.

"Laboratory Test Procedure for the New Car Assessment Program Side Impact Rigid Pole Test." Docket NHTSA- 2013-0052. 19 September 2013c.

National Highway Traffic Safety Administration. (2013d). NHTSA Announces Model Year 2014 Vehicles to be Rated Under Government 5-Star Safety Ratings Program [Press Release]. Retrieved from http://www.nhtsa.gov/About+NHTSA/Press+Releases/NHTSA+Announces+Model+Year+2014+Vehicles+to+be+Rated+Under+Government+5-Star+Safety+Ratings+Program.

"Opportunities Exist to Enhance NHTSA's New Car Assessment Program" GAO-05-370. April 2005.

Park, Brian T., et. al. "The Newly Enhanced U.S. NCAP: A First Look at Model Year 2011 Ratings." Paper 11-0440, 22nd International Conference on the Enhanced Safety of Vehicles, Washington, D.C. June 13-16, 2011.

"The New Car Assessment Program; Suggested Approaches for Enhancements." 72 Federal Register 16 (25 January 2007a), pp 3473-3477.

"The New Car Assessment Program Suggested Approaches for Future Program Enhancements." DOT HS 810 698. January 2007b.

APPENDIX

Table A1.
Frontal NCAP Driver Average Probability of Injury Readings and Number of Star Ratings by Model Year

	2011	2012	2013	2014	2015	2014/2015 Combined	All Years
Number of Ratings	64	69	57	42	24	66	256
Average Probability of Head Injury	0.005	0.005	0.006	0.008	0.006	0.008	0.006
Average Probability of Neck Injury	0.069	0.069	0.071	0.070	0.070	0.070	0.070
Average Probability of Chest Injury	0.029	0.032	0.021	0.026	0.024	0.025	0.027
Average Probability of Leg Injury	0.018	0.015	0.012	0.013	0.010	0.012	0.014
Average Occupant Combined Injury Probability	0.117	0.116	0.106	0.113	0.107	0.111	0.113
Average Star Rating	4.28	4.30	4.44	4.29	4.42	4.33	4.34

Table A2.
Frontal NCAP Driver Star Ratings by Model Year

	2 Star Rating		3 Star Rating		4 Star Rating		5 Star Rating		Total
Model Year	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	
2011	2	3.13%	8	12.50%	24	37.50%	30	46.88%	64
2012	2	2.90%	4	5.80%	34	49.28%	29	42.03%	69
2013	1	1.75%	5	8.77%	19	33.33%	32	56.14%	57
2014	1	2.38%	4	9.52%	19	45.24%	18	42.86%	42
2015	1	4.17%	0	0.00%	11	45.83%	12	50.00%	24
2014/2015 Combined	2	3.03%	4	6.06%	30	45.45%	30	45.45%	66
Total	**7**	**2.73%**	**21**	**8.20%**	**107**	**41.80%**	**121**	**47.27%**	**256**

Table A3.
Frontal NCAP Front Passenger Average Probability of Injury Readings and Number of Star Ratings by Model Year

	2011	2012	2013	2014	2015	2014/2015 Combined	All Years
Number of Ratings	64	69	57	42	24	66	256
Average Probability of Head Injury	0.015	0.013	0.012	0.008	0.011	0.009	0.012
Average Probability of Neck Injury	0.104	0.094	0.088	0.082	0.081	0.082	0.092
Average Probability of Chest Injury	0.018	0.016	0.014	0.014	0.018	0.015	0.016
Average Probability of Leg Injury	0.022	0.022	0.014	0.015	0.013	0.015	0.018
Average Occupant Combined Injury Probability	0.151	0.140	0.124	0.116	0.119	0.117	0.133
Average Star Rating	3.55	3.71	4.05	4.24	4.25	4.24	3.88

Table A4.
Frontal NCAP Front Passenger Star Ratings by Model Year

	2 Star Rating		3 Star Rating		4 Star Rating		5 Star Rating		Total
Model Year	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	
2011	12	18.75%	12	18.75%	33	51.56%	7	10.94%	64
2012	4	5.80%	21	30.43%	35	50.72%	9	13.04%	69
2013	2	3.51%	9	15.79%	30	52.63%	16	28.07%	57
2014	3	7.14%	2	4.76%	19	45.24%	18	42.86%	42
2015	1	4.17%	2	8.33%	11	45.83%	10	41.67%	24
2014/2015 Combined	4	6.06%	4	6.06%	30	45.45%	28	42.42%	66
Total	**22**	**8.59%**	**46**	**17.97%**	**128**	**50.00%**	**60**	**23.44%**	**256**

Table A5.
Side MDB NCAP Driver Average Probability of Injury Readings and Number of Star Ratings by Model Year

	2011	**2012**	**2013**	**2014**	**2015**	**2014/2015 Combined**	**All Years**
Number of Ratings	64	66	52	46	29	75	257
Average Probability of Head Injury	0.001	0.001	0.001	0.001	0.000	0.001	0.001
Average Probability of Chest Injury	0.070	0.059	0.040	0.040	0.040	0.040	0.052
Average Probability of Abdomen Injury	0.027	0.023	0.018	0.014	0.013	0.014	0.020
Average Probability of Pelvis Injury	0.006	0.005	0.004	0.003	0.003	0.003	0.004
Average Occupant Combined Injury Probability	0.099	0.085	0.062	0.058	0.055	0.057	0.076
Average Star Rating	4.44	4.56	4.81	4.87	4.86	4.87	4.67

Table A6.
Side MDB NCAP Driver Star Ratings by Model Year

	1 Star Rating		2 Star Rating		3 Star Rating		4 Star Rating		5 Star Rating		Total
Model Year	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	
2011	1	1.56%	3	4.69%	6	9.38%	11	17.19%	43	67.19%	64
2012	0	0.00%	2	3.03%	7	10.61%	9	13.64%	48	72.73%	66
2013	0	0.00%	0	0.00%	1	1.92%	8	15.38%	43	82.69%	52
2014	0	0.00%	0	0.00%	1	2.17%	4	8.70%	41	89.13%	46
2015	0	0.00%	0	0.00%	1	3.45%	2	6.90%	26	89.66%	29
2014/2015 Combined	0	0.00%	0	0.00%	2	2.67%	6	8.00%	67	89.33%	75
Total	**1**	**0.39%**	**5**	**1.95%**	**16**	**6.23%**	**34**	**13.23%**	**201**	**78.21%**	**257**

Table A7.
Side MDB NCAP Rear Passenger Average Probability of Injury Readings and Number of Star Ratings by Model Year

	2011	2012	2013	2014	2015	2014/2015 Combined	All Years
Number of Ratings	62	66	51	46	29	75	254
Average Probability of Head Injury	0.009	0.021	0.007	0.009	0.005	0.007	0.011
Average Probability of Pelvis Injury	0.079	0.073	0.041	0.044	0.033	0.040	0.058
Average Occupant Combined Injury Probability	0.087	0.092	0.047	0.052	0.038	0.047	0.068
Average Star Rating	4.37	4.38	4.86	4.80	5.00	4.88	4.62

Table A8.
Side MDB NCAP Rear Passenger Star Ratings by Model Year

	1 Star Rating		2 Star Rating		3 Star Rating		4 Star Rating		5 Star Rating		Total
Model Year	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	
2011	0	0.00%	7	11.29%	5	8.06%	8	12.90%	42	67.74%	62
2012	1	1.52%	6	9.09%	5	7.58%	9	13.64%	45	68.18%	66
2013	0	0.00%	1	1.96%	1	1.96%	2	3.92%	47	92.16%	51
2014	0	0.00%	3	6.52%	0	0.00%	0	0.00%	43	93.48%	46
2015	0	0.00%	0	0.00%	0	0.00%	0	0.00%	29	100.00%	29
2014/2015 Combined	0	0.00%	3	4.00%	0	0.00%	0	0.00%	72	96.00%	75
Total	**1**	**0.39%**	**17**	**6.69%**	**11**	**4.33%**	**19**	**7.48%**	**206**	**81.10%**	**254**

Table A9.
Side Pole NCAP Driver Average Probability of Injury Readings and Number of Star Ratings by Model Year

	2011	2012	2013	2014	2015	2014/2015 Combined	All Years
Number of Ratings	62	64	48	46	28	74	248
Average Probability of Head Injury	0.024	0.064	0.036	0.015	0.012	0.014	0.034
Average Probability of Pelvis Injury	0.117	0.082	0.064	0.048	0.054	0.050	0.078
Average Occupant Combined Injury Probability	0.137	0.136	0.098	0.062	0.065	0.063	0.107
Average Star Rating	4.10	4.23	4.52	4.85	4.82	4.84	4.44

Table A10.
Side Pole NCAP Driver Star Ratings by Model Year

	1 Star Rating		2 Star Rating		3 Star Rating		4 Star Rating		5 Star Rating		Total
Model Year	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	**Count**	**% Total**	
2011	5	8.06%	7	11.29%	2	3.23%	11	17.74%	37	59.68%	62
2012	4	6.25%	5	7.81%	6	9.38%	6	9.38%	43	67.19%	64
2013	1	2.08%	1	2.08%	3	6.25%	10	20.83%	33	68.75%	48
2014	0	0.00%	0	0.00%	1	2.17%	5	10.87%	40	86.96%	46
2015	0	0.00%	0	0.00%	1	3.57%	3	10.71%	24	85.71%	28
2014/2015 Combined	0	0.00%	0	0.00%	2	2.70%	8	10.81%	64	86.49%	74
Total	**10**	**4.03%**	**13**	**5.24%**	**13**	**5.24%**	**35**	**14.11%**	**177**	**71.37%**	**248**

For the star rating tables, percentages may not total 100 due to rounding.

EVALUATION of MOTORCYCLE HELMET USABILITY and STABILITY in the CONSUMER SAFETY RATING PROGRAM 'CRASH'

Andrew McIntosh
McIntosh Consultancy and Research
Monash Injury Research Institute, Monash University
Australia

Basuki Suratno
Transport for NSW
Australia

Jack Haley
NRMA Motoring and Services
Australia

Jessica Troung
Transport Accident Commission
Australia

Paper Number 15-0292

ABSTRACT

A review of certain aspects of the Consumer Rating and Assessment of Safety of Helmets (CRASH) for motorcyclists was undertaken. The paper examines the relationships between the assessment of helmet stability with volunteers and other usability assessments. The paper reports on how these assessments are incorporated into the Consumer Rating and Assessment of Safety Helmets (CRASH) and general relationships between Safety Scores and Ergonomic Scores. Ninety (90) motorcycle helmets were evaluated in the years 2011 to 2014 involving dynamic stability tests, dynamic strength of retention tests and usability tests with six participants. All helmets complied with AS/NZS 1698: 2006. The participants rated each helmet across ten items using a five point Likert scale. Forward, rearward and lateral pull tests were performed on each participant with each helmet. The force required to move the helmet with respect to the scalp was measured. The analysis revealed a number of important findings. First, safety performance tests are not correlated with ergonomic assessments of the helmets, including formal usability assessments. This observations highlights the importance of providing both safety and ergonomic information to motorcyclists. Helmet mass ranged between approximately 1 kg and 2 kg in the sample assessed. On the assumption that the total Safety Score reflects a helmet that offers greater protection in a crash, a heavier helmet within the sample assessed offers more protection to the motorcyclist. Full face helmet types also performed better on total Safety Score than the open face helmet. Full face helmets were heavier than open face styled helmets. Differences in the total Ergonomic Score by helmet type were fewer than those observed with Safety Scores. There were strong correlations between the rater responses between pairs of questions regarding comfort, fit, but not restraint adjusment. Although there is some overlap between these questions, each question appears to elicit a slightly different response across all helmet types and raters. Helmet mass is either not correlated or weak to moderately correlated with user ratings. This suggests that the raters are considering other factors, not simply mass, when rating helmet weight. The raters might be considering the mass distribution, for example. In general, helmet stability as measured quasi-statically on each rater was weakly associated with the raters' assessment of the helmet. This suggests that the motorcylist's impression of fit is not a strong indicator of helmet stability. Ease of use of operation was only weakly associated, based on these results, with the forces required to displace the helmet on the rater's head. Correlations between the stability test forces by direction (front, rear and lateral) were strong.

INTRODUCTION

Powered two wheeler (PTW), including motorcycle, safety is a major global issue. Although the incidence rates for motorcyclist injury and death have decreased over decades, the absolute number of cases has been increasing as a result of increased motorcyclist exposure.[1] Motorcycle helmets are one method for preventing head injury in motorcycle crashes.[2,3] Geographical regions and countries have unique systems in place that regulate the supply and use of motorcycle helmets. A central element of these systems are motorcycle helmet standards. Standards are referred to with regards to both sale and on-road use. Standards are critical in terms of providing motorcyclists with effective helmets. However, Standards do not assess all aspects of helmet use or performance.[4,5] Consumer rating programs can address a range of usability and performance issues that affect motorcyclist safety.[5] The paper assists in identifying how usability assessments can be applied in consumer information programs and what factors may be important in terms of helmet stability at the point of sale.

The 2011 IRTAD report highlighted that despite reductions in mortality rates for PTW operators, the relative risk remains much greater for PTW groups than passenger car occupants; in European countries the relative rate is between 17 and 20 times in some countries.[1] In the USA motorcycle fatalities comprised 14% of the total road fatalities in 2010. [1] In other regions, PTW fatalities represent the majority of road fatalities, e.g. 71% in Cambodia and 59% in Malaysia. [1] Helmet use has been shown to be effective in preventing head injury. A 2008 Cochrane review of motorcycle helmet estimated that helmets reduce the risk of head injury by 69% and death by 42%. [2] A study of trauma centre admissions found that helmets were associated with a significant reduction in intracranial injury likelihood of 66% amongst motorcyclists.[6] Therefore, an important element of a road safety system is helmet use.[1-6]

Helmets need to fit well and be comfortable.[7] Maintaining the positional stability of a motorcycle helmet in general use and a crash is a fundamental performance requirement. [7-10] Helmet stability is assessed in many helmet standards, but ergonomic surveys of motorcyclists have shown that helmet design and the user influence helmet stability.[4,7] A survey of 216 motorcyclists found that "the size of the in-use motorcycle helmets did not correspond well to the predicted size based on head dimensions, although motorcyclists were generally satisfied with comfort and fit."[7] The forces required to displace each wearer's helmet were also observed to be low, around 25 N. [7] The paper will examine the relationships between the assessment of helmet stability with volunteers and other usability assessments. The paper will report on how these assessments are incorporated into the Consumer Rating and Assessment of Safety Helmets (CRASH) and general relationships between Safety Scores and Ergonomic Scores.

METHODS

Ninety (90) motorcycle helmets were evaluated in the years 2011 to 2014 involving a range of tests, including formal usability tests with six participants (raters). All helmets complied with AS/NZS 1698: 2006. The helmet sample was derived from advice from wholesalers, retailers and consumers, as well as historical trends and coverage of specific categories of helmet types. Over three CRASH rating periods Dual Sport, Full Face, Flip Up, Motocross, Open Face, and Open Face with Visor helmet types have been tested.

The test program methods were reported in 2013.[5] In short the total Safety Score reflects performance on: helmet coverage; dynamic stability; high level impact energy attenuation (2.5 m drop); low level impact energy attenuation (0.8 m drop); kerb anvil high level impact energy attenuation (2.5 m drop); and dynamic retention strength. The total Ergonomic Score reflects usability rater assessment and performance on visor fogging, splash resistance, aerodynamics, in-helmet noise, mass and field of view. The rater assessment comprises (i) a standard protocol with ten questions and (ii) in-situ force to commence helmet displacement measured. All participants had a head circumference equivalent to the ISO "J" headform (57-58 cm). The participants rated each helmet across ten items using a five point Likert scale. Forward, rearward and lateral pull tests were performed on each participant with each helmet.[5] The force required to move the helmet with respect to the scalp was measured.[5,7]

The following rating questions were considered in this analysis: Question 1. Does the helmet allocated to you fit your head comfortably; Question 2 From no pain or pressure points (score 5) to very uncomfortable (score 1) please rate the helmet with reference to your face, chin and head; Question 3 Please rate the weight of the helmet in terms

of its comfort from very comfortable (5) to very uncomfortable; Question 4 Please rate the helmet in terms of comfort and fit overall from very good (5) to very poor (1); and Question 10 How easy is it to adjust and tighten the chin strap or restraint system from very easy (5) to very difficult (1). Questions 5 to 9 were not considered relevant to the assessment as they dealt with other helmet features, e.g. the visor operation. The higher the score on the Likert scale the better the rating. Data were collated and analysed using the following statistical methods. Descriptive statistics for the total Safety and Ergonomic Scores and mass were prepared. Correlations between the total scores and mass were assessed, as were differenes in the mean total scores and mass by helmet type. Bivariate correlations between rater scores (1 to 5) and helmet mass, and pull forces were assessed.

RESULTS

Over three CRASH rating periods a total of ninety helmets were evaluated for safety and ergonomic performance. The total Safety and Ergonomic Scores and helmet masses are presented in figures 1 to 3. The helmets are de-identified. Helmets have been sorted by type (DS Dual Sport; FF Full Face; FU Flip Up; MC Motocross; OF Open Face; OFV Open Face with Visor). There were four Dual Sport helmets (4.4%), 51 Open Face (56.7%), 10 Flip Up (11.1%), 2 Motocross (2.2%), 12 Open Face (13.3%) and 11 Open Face with Visor helmets (12.2%). Table 1 presents descriptive statistics for total safety and ergonomic scores and helmet mass by helmet type. There are significant differences ($p<0.05$) in the mean helmet mass by type: DS, FF,FU and MC are heavier than OF and OFV; OFV is heavier than OF; and, FU is heavier than FF. There are significant differences ($p<0.05$) in the mean helmet total Safety Scores by type: DS, FF and FU have higher mean Safety Scores than OF. There are significant differences ($p<0.05$) in the mean helmet Ergonomic Scores by type: FF has a greater mean than OFV. Figure 4 shows that in terms of the total Safety and Ergonomic Scores, there is a strong correlation between total Safety Score ($r^2 = 0.275$) and mass and none between total Ergonomic Score and mass. Therefore, as helmet mass increases there is an increase in the safety rating as scored in the CRASH program. The analysis of total Ergonomic Score with respect to mass is confounded by the components in the total score. Total Ergonomic and Safety Scores were not correlated.

Table 1.
Descriptive statistics for helmet total safety and ergonomic scores and mass by helmet type (n=90).

		Count	Mean	Median	Standard Deviation	Minimum	Maximum
DS	Mass (kg)	4	1.68	1.69	0.12	1.53	1.8
	Safety	4	62	61	8	54	72
	Ergonomic	4	54	53	10	44	68
FF	Mass (kg)	51	1.61	1.62	0.1	1.36	1.94
	Safety	51	58	59	11	35	76
	Ergonomic	51	57	55	10	43	77
FU	Mass (kg)	10	1.78	1.78	0.1	1.58	1.96
	Safety	10	57	58	13	35	74
	Ergonomic	10	49	49	8	32	62
MC	Mass (kg)	2	1.66	1.66	0.01	1.65	1.67
	Safety	2	59	59	17	47	71
	Ergonomic	2	44	44	2	42	45
OF	Mass (kg)	12	1.18	1.17	0.15	0.96	1.43
	Safety	12	38	36	12	21	58
	Ergonomic	12	50	49	8	37	62
OFV	Mass (kg)	11	1.34	1.34	0.11	1.17	1.51
	Safety	11	51	49	10	37	66
	Ergonomic	11	46	47	5	36	55

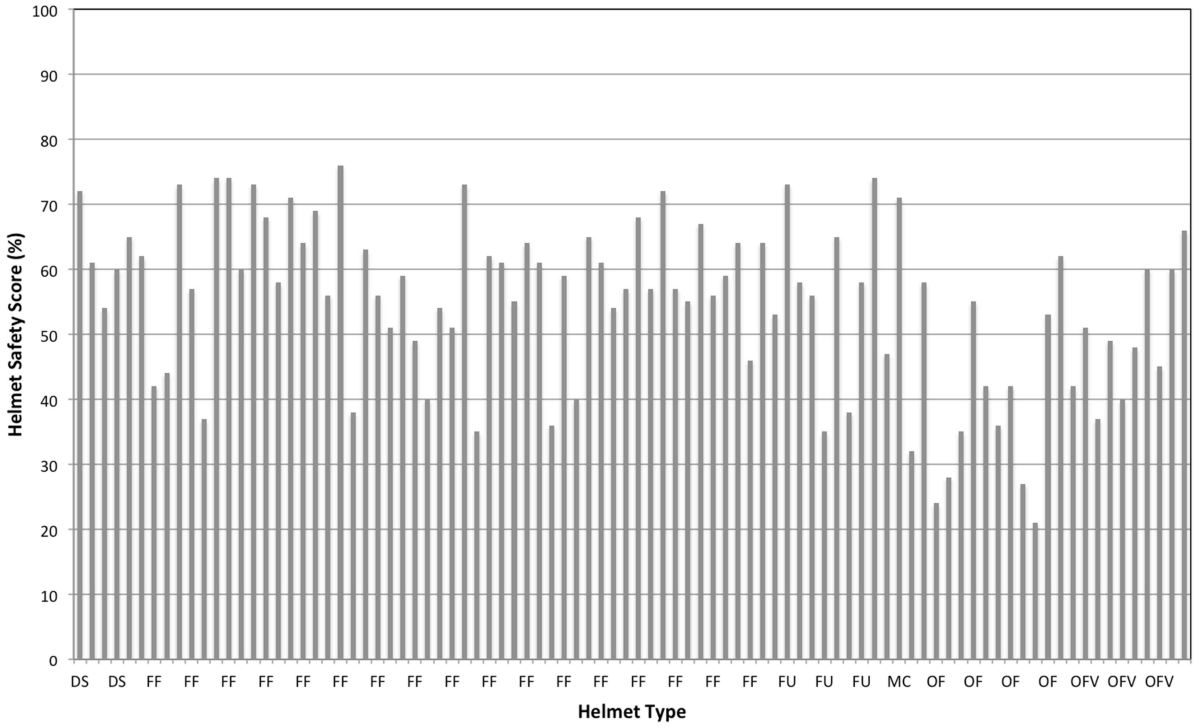


***Figure 1*. Final total safety score by helmet model grouped under type.**

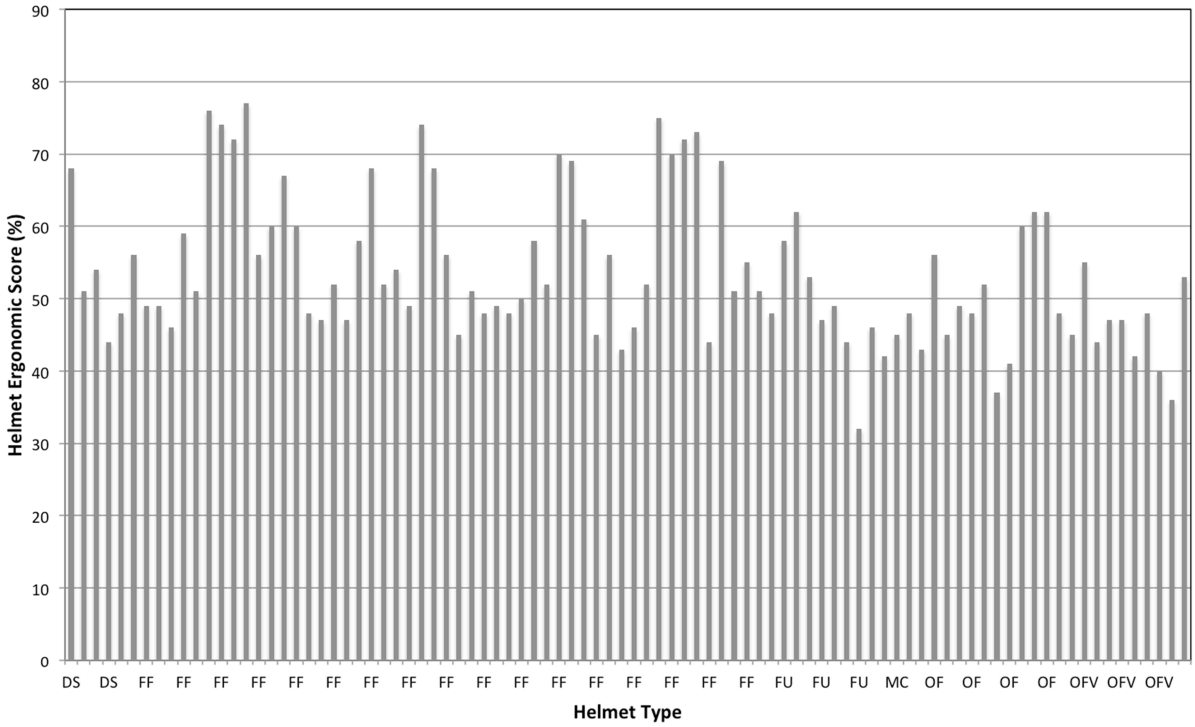


***Figure 2*. Final total ergonomic score by helmet model grouped under type.**

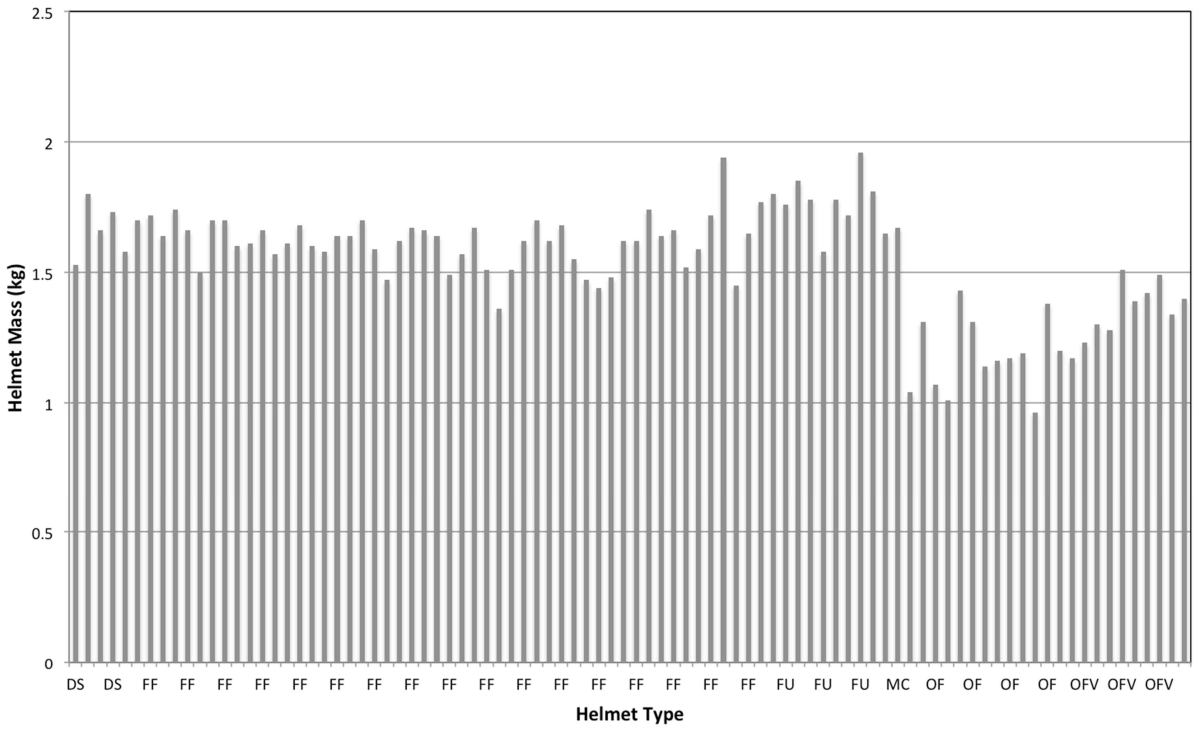


Figure 3. **Helmet mass by helmet model grouped under type.**

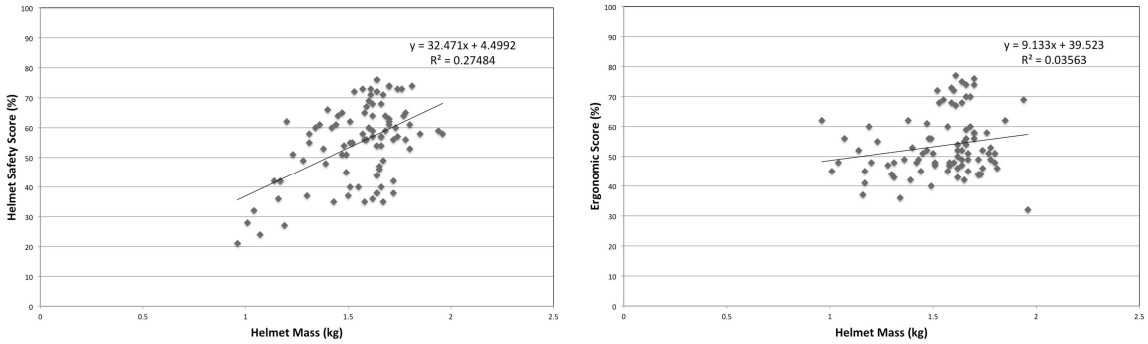


Figure 4. **Helmet mass by total Safety Score (left) and Ergonomic Score (right).**

The component ergonomic scores were examined for relationships between individual ratings on specific questions, helmet mass and stability as measured quasi-statically on the rater (table 2). A number of significant correlations were observed. Analyses were not adjusted for potentially confounding factors, e.g. helmet type, rater experience and preference.

Table 2.
Correlations between individual rater scores, helmet mass and forces measured in individual stability tests. The average force represents the average of three tests in each pull direction. PC is Pearson Correlation. Significance tests are two-tailed. ** Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).

		Q1	Q2	Q3	Q4	Q10	Mass (kg)	Rear (N)	For. (N)	Lat. (N)
Q1	PC	1	.730**	.522**	.789**	.226**	-0.02	-.172**	-.159**	-.097*
	Sig.		0	0	0	0	0.655	0	0	0.028
	N	509	509	509	509	509	509	509	509	509
Q2	PC	.730**	1	.607**	.811**	.121**	-.136**	-.208**	-.132**	-0.071
	Sig.	0		0	0	0.006	0.002	0	0.003	0.11
	N	509	509	509	509	509	509	509	509	509
Q3	PC	.522**	.607**	1	.554**	0.08	-.243**	-.219**	-.206**	-.152**
	Sig.	0	0		0	0.072	0	0	0	0.001
	N	509	509	509	509	509	509	509	509	509
Q4	PC	.789**	.811**	.554**	1	.214**	-0.077	-.155**	-.132**	-0.055
	Sig.	0	0	0		0	0.083	0	0.003	0.214
	N	509	509	509	509	509	509	509	509	509
Q10	PC	.226**	.121**	0.08	.214**	1	.215**	.172**	0.065	.100*
	Sig.	0	0.006	0.072	0		0	0	0.144	0.024
	N	509	509	509	509	509	509	509	509	509
Helmet Mass (kg)	PC	-0.02	-.136**	-.243**	-0.077	.215**	1	.443**	.350**	.352**
	Sig.	0.655	0.002	0	0.083	0		0	0	0
	N	509	509	509	509	509	509	509	509	509
Average Rear Pull Force (N)	PC	-.172**	-.208**	-.219**	-.155**	.172**	.443**	1	.643**	.616**
	Sig.	0	0	0	0	0	0		0	0
	N	509	509	509	509	509	509	509	509	509
Average Forward Pull Force (N)	PC	-.159**	-.132**	-.206**	-.132**	0.065	.350**	.643**	1	.778**
	Sig.	0	0.003	0	0.003	0.144	0	0		0
	N	509	509	509	509	509	509	509	509	509
Average Lateral Pull Force (N)	PC	-.097*	-0.071	-.152**	-0.055	.100*	.352**	.616**	.778**	1
	Sig.	0.028	0.11	0.001	0.214	0.024	0	0	0	
	N	509	509	509	509	509	509	509	509	509

DISCUSSION

The analysis of ninety motorcycle helmets revealed a number of important findings. First, safety performance tests are not correlated with ergonomic assessments of the helmets, including formal usability assessments. This observations highlights the importance of providing both safety and ergonomic information to motorcyclists to assist in purchasing decisions that might influence long term helmet use. This underpins the rationale for the CRASH program. On one hand the helmet purchaser may select a helmet that is comfortable, for example, but this does not provide an indication of the helmet's potential safety performance in a crash. On the other hand, a focus on safety performance alone does not identify areas where improvements might be made to encourage helmet use and selection. Helmet mass was correlated with the total Safety Score. On the assumption that the total Safety Score reflects a helmet that offers greater protection in a crash, a heavier helmet within the sample assessed offers more protection to the motorcyclist. Helmet mass ranged between approximately 1 kg and 2 kg in the sample assessed. Full face helmets were heavier than open face styled helmets. Full face helmet types (DS, FU and FF) also performed better on total Safety Score than the open face helmet. Differences in the total Ergonomic Score by helmet type were fewer than those observed with Safety Scores. These analyses are very general, because they do not consider the individual elements that make up the total scores. For example, the total Ergonomic Score comprises aerodynamic and in helmet noise assessments that tend to favour full face type helmets and are less influenced by helmet mass.

The analysis of comfort and fit related usability test questions, helmet mass and user stability tests showed a number of significant correlations (table 2). There were generally strong correlations between the rater responses between pairs of questions 1, 2, 3 and 4. Correlations between questions 1 to 4 and 10 were weak. Further analyses are required to assess rater and helmet type effects on these associations. Although there is some overlap between these questions, each question appears to elicit a slightly different response across all helmet types and raters. Helmet mass is either not correlated or weak to moderately correlated with user ratings. Question 3 is a direct rating of helmet weight; there is a significant but weak association between Question 3 and measured helmet mass. This suggests that the users are considering other factors, not simply mass, in Question 3. The raters might be considering the mass distribution, for example. The experiences of each rater might also influence the assessment, e.g. a rater accustomed to wearing a 'heavy' helmet, might be comfortable with helmets in that mass range. Further analysis is required.

In general, helmet stability as measured quasi-statically on each rater was weakly associated with the raters' responses to questions 1, 2, 3, 4 and 10 (table 2). Therefore, the rater's assessment of fit and comfort are not subsitutes for stability as assessed in these tests. This requires further analysis. It does suggest that the motorcylist's impression of fit is not a strong indicator of helmet stability. Therefore, a motorcyclist might purchase a helmet based on fit that does not offer the level of stability that they expect. Question 10 relates to the operation of the restraint system. Ease of use of operation was only weakly associated, based on these results, with the forces required to displace the helmet on the rater's head.

Correlations between the stability test forces by direction were strong. Further research is required to account for potential confounding factors, e.g. helmet type and head dimensions.[7]

CONCLUSIONS

Research has identified the importance of helmet stability and retention in terms of usability and crash performance. An analysis of the CRASH helmet data provides a greater understanding of the user's perceptions of the helmet function and what needs to be assessed in a consumer safety rating program. This information can complement impact performance and standards compliance information to assist consumer decisions and helmet suppliers in improving helmet design.

The CRASH program is funded by Transport for NSW, NRMA Motoring and Services and Transport Accident Commission (Australia).

The opinions presented in this paper are those of the authors and do not represent those of the organisations they represent.

REFERENCES

[1] IRTAD. Road Safety Annual Report 2011: International Traffic Safety Data and Analysis Group IRTAD, OECD/ITF; 2012
[2] Liu BC, Ivers R, Norton R et al, (2008) Helmets for preventing injury in motorcycle riders, Cochrane Database Syst Rev. Jan 23;(1):CD004333
[3] Lin M-R & Krauss JF (2008) Methodological issues in motorcycle injury epidemiology; 40: 1653-1660
[4] McIntosh AS & Grzebieta R, (2013) Motorcycle helmet standards – harmonsation and specialisation ? 2013, 23rd International Technical Conference on the Enhanced Safety of Vehicles, May 27-30, Seoul, Paper Number 13-0160
[5] McIntosh AS, Suratno B, Haley J, Truoung J, (2013) Consumer rating and assessment of safety helmets for motorcyclists, 2013, 23rd International Technical Conference on the Enhanced Safety of Vehicles, May 27-30, Seoul, Paper Number 13-0157
[6] McIntosh AS, Curtis K, Rankin T, Cox M, Pang TY, McCrory P and Finch CF, (2013) Helmets prevent brain injuries in injured pedal- and motor-cyclists: A case series analysis of trauma centre presentations, Australian College of Road Safety Journal, 24: 11-20.
[7] Thai KT, McIntosh AS, Pang TY, Factors affecting motorcycle helmet use: Size selection, stability and position. Traffic Injury Prevention 2015 Apr 3;16(3):276-82. doi: 10.1080/15389588.2014.934366. Epub 2014 Nov 14
[8]. McIntosh AS & Lai A (2013) Motorcycle Helmets: Head and Neck Dynamics in Helmeted and Unhelmeted Oblique Impacts, Traffic Injury Prevention, 14:8, 835-844
[9] Hurt HH, Thom DR, Ouellet JV. (1998) Testing the positional stability of motorcycle helmets. Paper presented at: 16th International Technical Conference on the Enhanced Safety of Vehicles (ESV); 1998; Windsor, ON, Canada.
[10] Chinn B, Canaple B, Derler S, et al. (2001) COST 327: Motorcycle Safety Helmets. Brussels, Belgium: European Commission, Directorate General for Energy and Transport; 2001

EURO NCAP'S FIRST STEP TO ASSESS AUTONOMOUS EMERGENCY BRAKING (AEB) FOR VULNERABLE ROAD USERS

Richard Schram
Aled Williams
Michiel van Ratingen
European New Car Assessment Programme,
Belgium
on behalf of the Euro NCAP P-NCAP Working Group

Stephan Ryrberg (Chairman of ACEA TF NCAP Expert Group AEB)
Volvo Car Corporation
Sweden
Raimondo Sferco (Chairman of ACEA TF NCAP)
Ford Motor Company
Germany
on behalf of the ACEA, JAMA and KAMA joint WG on AEB

Paper Number: 15-0277

ABSTRACT
Following the implementation of AEB City and AEB Inter-Urban systems in Euro NCAP's safety rating in 2014, a third type of AEB technology, Autonomous Emergency Braking for Vulnerable Road Users (AEB VRU), will be added to the overall assessment of new vehicles in 2016. The introduction of AEB VRU will be done in two phases where in 2016 AEB Pedestrian is implemented followed by AEB Cyclists in 2018.

AEB VRU will be awarded as part of the assessment of Pedestrian Protection and represents the next step to improve the protection of vulnerable road users, complimentary to the existing subsystem tests to the vehicle front end. Following system tests in common pedestrian accident scenarios, more challenging and demanding cyclist scenarios are planned in a subsequent phase.

In close corporation with the car industry represented by the ACEA, JAMA and KAMA associations, Euro NCAP has developed detailed test and assessment procedures for AEB Pedestrian. The procedures are based on the existing car to car AEB test and assessment protocols and validated and checked for repeatability and reproducibility at several Euro NCAP laboratories. This paper describes both the test and assessment protocols.

BACKGROUND
In 2009, Euro NCAP introduced the overall rating scheme, which allows new technologies to be implemented in the safety assessment of a new vehicle. The new rating scheme consists of four areas of assessment, also called boxes, which together result in a single overall safety rating. The four areas of assessment are Adult Occupant Protection (AOP), Child Occupant Protection (COP), Pedestrian Protection (PP) and Safety Assist (SA).

Over the last few years, Euro NCAP has rapidly raised the requirements for better protection of vulnerable road users in the event of a crash which was seen to lag behind. After significant updates to the test and assessment protocols in the area of passive safety, the next logical step was to include the assessment of AEB systems. Due to the nature of pedestrian accidents with passenger cars, most AEB systems are only capable of mitigating these crashes and therefore, Euro NCAP considers these systems complementary to passive safety measures already in place.

At the moment different AEB Pedestrian systems are already available on the market, based on radar, (stereo) camera or a combination of these sensors. The performance of these systems vary significantly based on the sensors used. The first generation of camera based systems typically switch off during low ambient lighting conditions as classification of pedestrians in darkness is not reliable enough. Euro NCAP has based its requirements for AEB Pedestrian systems on best practice to push for further development of these lifesaving technologies.

WORKING GROUP

As for all Euro NCAP protocols, the development of the test procedure and assessment criteria was done within a collaborative Working Group. The Primary NCAP working Group, P-NCAP TWG in short, involved Euro NCAP members and laboratories and was given the task to deliver AEB Pedestrian protocols by the end of 2014, for implementation in 2016. Although car makers and suppliers were not directly involved in the working group, several meetings were organised between representatives of both sides to review the procedures. More importantly, the work of the group took advantage of and brought together the results delivered by several main research initiatives in Europe that where looking into the development of AEB test and assessment procedures.

Initiatives

In Europe, several initiatives have been running in parallel, all with the same goal of developing test procedures for assessing AEB systems in general and AEB Pedestrian systems in particular: "AEB", "AsPeCSS" and "vFSS". Besides this, ACEA started an internal project to develop so called *articulated* pedestrian targets.

The RCAR Autonomous Emergency Braking group, led by Thatcham aims at designing and implementing a testing and rating procedure for Autonomous Emergency Braking ("AEB") systems reflecting real world accident data. It is hoped that this will encourage the development of AEB systems that can avoid or mitigate the effects of car-to-pedestrian and car-to-car collisions seen in the most common crash types. The group mainly consists of insurance institutes, supported by Volvo Car Corporation, Subaru, Daimler and first-tier supplier Continental. The work of the RCAR AEB group is ongoing.

The European Commission sponsored FP7 project AsPeCSS (Assessment methodologies for forward looking Integrated Pedestrian and further extension to Cyclist Safety Systems) led by IDIADA had specific project goals to develop harmonised and standardised procedures for the assessment of forward looking integrated pedestrian safety systems. The project partners consisted of nine research institutes, three of which were Euro NCAP laboratories: BASt, IDIADA and TNO. From industry side, BMW, PSA and Toyota participated as car manufacturers and Autoliv, Bosch and TRW as first-tier suppliers. The AsPeCSS project ran from 2011 until 2014.

The third initiative was vFSS (Advanced Forward-Looking Safety Systems) led by DEKRA. The project was supported by the BASt as a Euro NCAP Test Lab. The Insurance Companies Allianz, GDV and KTI as well as the AUDI, BMW, Daimler, Ford, Honda, Opel, Porsche, Toyota and VW as industry partners. One of the main deliverable were the static pedestrian targets that were adopted in the Euro NCAP protocols. They will be described in further detail within this paper.

The outcome and deliverables of all the initiatives were extensively discussed within the working group and formed the basis for the decision on test scenarios and targets used.

TEST SCENARIOS AND TARGETS

There was a large overlap of the proposed test scenarios from the different initiatives, based on an extensive analysis of real world car-to-pedestrian accidents mainly from Germany, Great-Britain and France. Overlaying the proposed test scenarios, the P-NCAP TWG agreed to focus on four test scenarios for AEB Pedestrian. In the Car-to-VRU Farside Adult (CVFA) scenario, the running pedestrian crosses the vehicle path from the farside. It represents the situation where a pedestrian first crosses another lane before entering the lane in which the vehicle is driving. The timing is set such that without any AEB reaction from the vehicle, the adult pedestrian target would end up contacting the centreline of the vehicle. The second adult pedestrian scenario represents a pedestrian stepping of the sidewalk into the lane the vehicle is driving in. There are two variants with different impact locations of 25 and 75% of the vehicles width (CVNA-25 & CVNA-75).
The final scenario is the most challenging one, where a child pedestrian appears from behind two parked vehicles and directly stepping into the lane the vehicle is in. Similar to the farside adult scenario, the impact location will be in the middle of the vehicle (CVNC).

Car-to-VRU Farside Adult (CVFA)

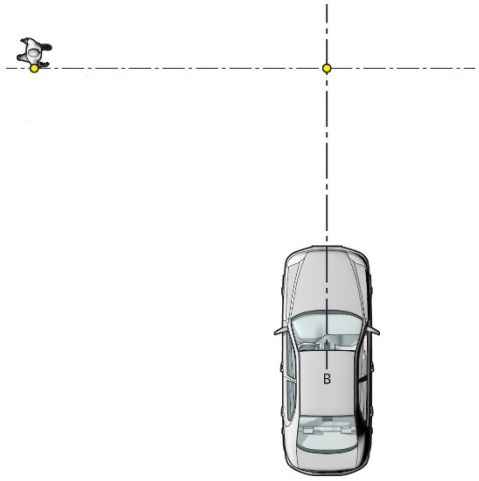


Figure 1. AEB Pedestrian scenario, CVFA

Car-to-VRU Nearside Adult (CVNA-25 & CVNA-75)

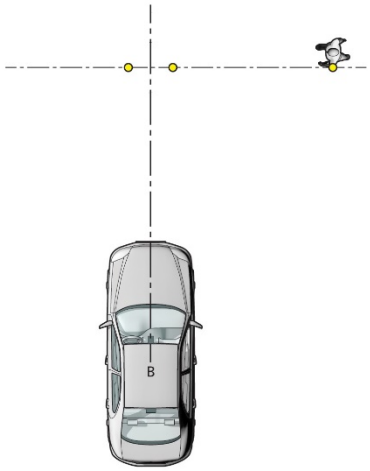


Figure 2. AEB Pedestrian scenario, CVNA-25 & 75

Car-to-VRU Nearside Child (CVNC)

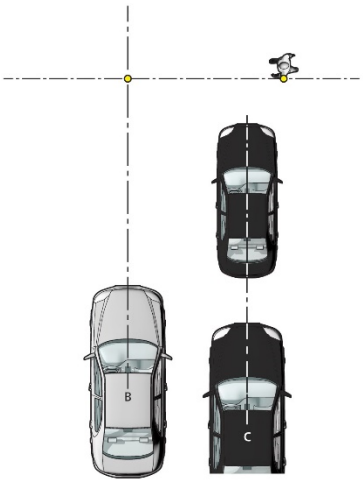


Figure 3. AEB Pedestrian scenario, CVNC

Figure 4. Euro NCAP Pedestrian Targets (EPTa & EPTc)

For the all of the above described AEB Pedestrian scenarios, only the AEB function (i.e. the vehicle braking without driver involvement) is tested and assessed. Possible Forward Collision Warning (FCW) functionality (the driver responding to a warning by applying the brakes) is not taken into account as it assumed that in this type of situations, there is generally not sufficient time to react to a warning.

Incremental Speed Approach

All of the scenarios will be tested with an incremental approach. Starting at a low speed of 20 km/h, the approach speed of the Vehicle Under Test (VUT) is stepwise increased by 5 km/h up to a maximum test speed of 60 km/h. For each run, the vehicle's speed reduction is recorded.

Euro NCAP Pedestrian Targets

A large number of different pedestrian targets have been developed over time. Following several workshops, it was decided to take the pedestrian targets as developed within vFSS as a basis. Further events to verify the radar cross section and infrared reflectivity were performed and it was confirmed that the static dummy was detectable by both radar and camera based systems. The figures 5 and 6 show the stance and statures of the adult and child pedestrian targets.

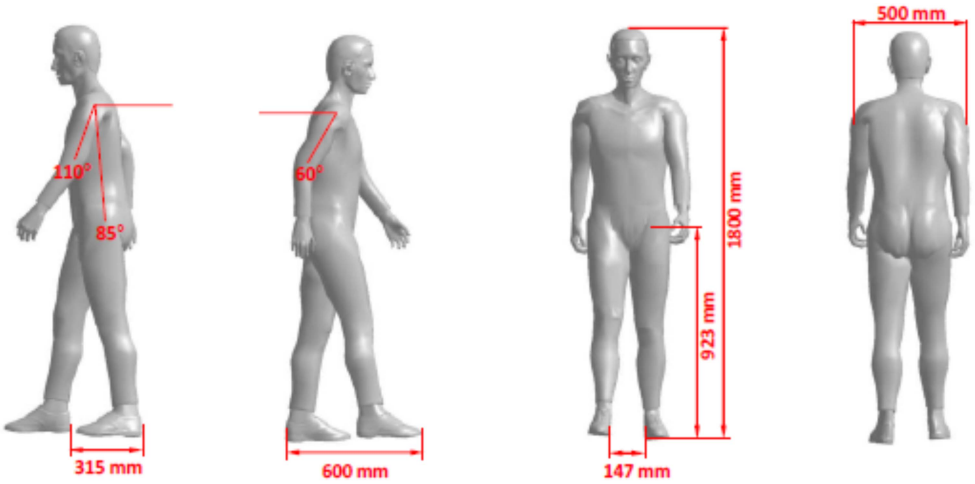


Figure 5. Euro NCAP adult Pedestrian Target stance and dimensions

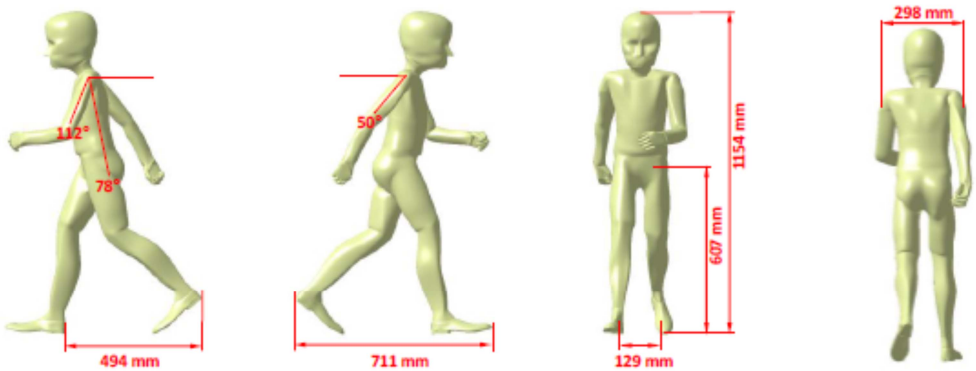


Figure 6. Euro NCAP child Pedestrian Target stance and dimensions

At a later stage, industry started the development of an articulated dummy with moving legs that would reflect a more realistic Doppler image, used by more advanced radar systems. Apart from the moving legs, the dummy radar cross section, clothing and stature are the same as for the static pedestrian targets. It was agreed by Euro NCAP to consider specifying this dummy as the pedestrian target in the protocols when all parties involved had verified and approved it. The evaluation and fine-tuning of this dummy is ongoing at the time of writing of this paper.

Test equipment and test track

Euro NCAP uses different facilities for all of its tests. To ensure repeatable and reproducible results now and in the future, the WG had already decided to set strict tolerances for testing all of the AEB systems, even though it was acknowledged that this may not always be necessary to evaluate the performance of these systems in the scenarios described earlier. The tolerances used for the AEB Pedestrian tests are listed below:

- Speed of VUT (GPS-speed)	Test speed + 0.5 km/h
- Lateral deviation from test path	0 ± 0.05 m
- Yaw velocity	0 ± 1.0 °/s
- Steering wheel velocity	0 ± 15.0 °/s
- Speed of EPT during steady state	
o CVFA	8 ± 0.2 km/h
o CVNA	5 ± 0.2 km/h
o CVNC	5 ± 0.2 km/h
- EPT Steady state	
o Nearside	3.0 m from vehicle centerline
o Farside	4.5 m from vehicle centerline

Due to these strict tolerances, all of the Euro NCAP test facilities are using both steering and brake robots to control the vehicle during test.

Another, less controllable, influencing factor is weather condition. The tracks used for the assessment are spread over Europe with different climates. Although the weather may influence the performance of the systems, it is thought that in day-to-day use these systems also encounter various weather conditions and hence should be robust enough to deal with normal variations. To minimise test to test variability, limits are set to temperature (between 5 and 40°C) and wind (below 10 m/s). There may be no precipitation falling and horizontal visibility at ground level must be greater than 1km. Finally, the natural ambient illumination must be homogenous in the test area and in excess of 2000 lux for daylight testing. For practical reasons, Euro NCAP does not physically test the performance of these systems in low ambient lighting conditions even though accident data reveals this is relevant. It is also ensured that testing is not performed driving towards or away from the sun when there is direct sunlight.

ASSESSMENT

The assessment of AEB Pedestrian systems is based on two aspects: the Autonomous Emergency Braking function, i.e. how well the system reacts to an imminent pedestrian impact, and the Human Machine Interface. The latter assessment carries less weight but is important to promote the use of the system.

Assessment of AEB function

The only assessment criterion used is the impact speed reduction. For each run into the target at incremental speed, a full score is given when the impact is completely avoided. Where a contact occurs, the points are awarded on a sliding scale basis for speeds up to and including 40 km/h, where the proportion of speed reduction based on the relative test speed determines the proportion of available points scored.

$$Score_{\ test\ speed} = ((Vtest - Vimpact)/Vtest) \times points_{test\ speed}$$

For test speeds above 40km/h points are available on a pass/fail basis only. For each of these test speeds points are awarded when a speed reduction of at least 20 km/h is achieved related to actual test speed. It was acknowledged within the working group that at the moment it is not realistic to ask for full avoidance at speeds above 40 km/h as this would require a relatively early AEB activation at a point in time where the pedestrian has not yet entered the vehicle path. By requiring the impact speed to be reduced with at least 20 km/h, it brings the severity of the crash into the range where passive safety measures are designed to work.

The number of points available for the different test speeds is the same for each of the 4 scenarios based on exposure multiplied by injury levels coming from the accident data. The available point distributions for all scenarios are shown in the figure below.

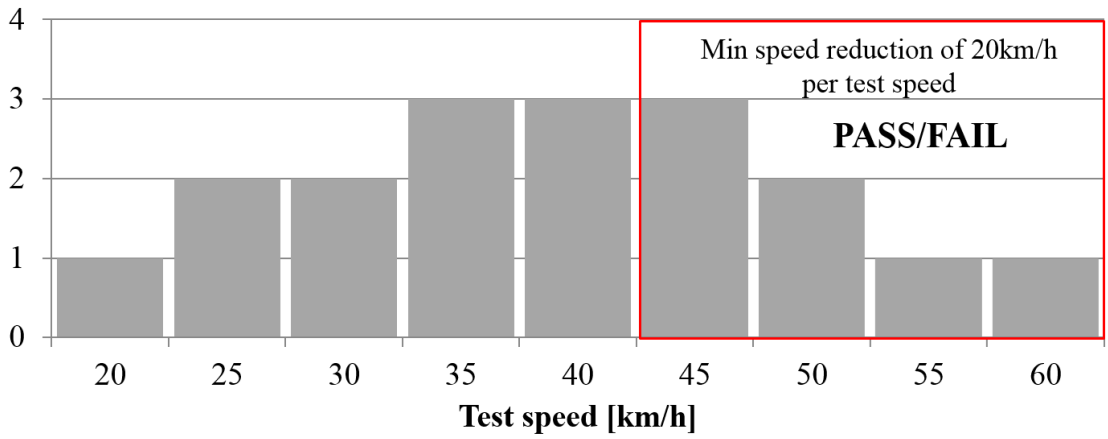


Figure 6. Maximum points per test speed

Human Machine Interface

The effectiveness of the whole system does not only depend on the speed reduction achieved. The ON/OFF rate of the system is highly influencing the actual performance. At this moment, Euro NCAP has not defined qualitative criteria for warning due to the limited knowledge available on this subject. However, points are

awarded to systems that encourage use, for instance making it less easy to switch the system off or avoiding the system to automatically switch off at low ambient lighting conditions.

Scoring
For AEB Pedestrian, only the autonomous emergency braking functionality is considered and HMI points will only be awarded if the AEB system is default ON at the start of every journey.

For the total score of these systems, the normalized sub-scores (as a percentage of the maximum points available) of AEB and HMI are weighted and summed.

$$AEB\ Pedestrian\ total\ score = (AEB\ score\ x\ 5) + (HMI\ score\ x\ 1)$$

A scoring example for an AEB Pedestrian system is provided below.

Table 1
Example of AEB function test results in CVFA scenario

Vtest	$points_{test\ speed}$	Vimpact	$Score_{test\ speed}$
20 km/h	1.000	0 km/h	1.000
25 km/h	2.000	0 km/h	2.000
30 km/h	2.000	0 km/h	2.000
35 km/h	3.000	0 km/h	3.000
40 km/h	3.000	20 km/h	1.500
45 km/h	3.000	25 km/h	3.000
50 km/h	2.000	30 km/h	2.000
55 km/h	1.000	40 km/h	0.000
60 km/h	1.000	Not tested	0.000
Total	**18.000**		**14.500**
Normalised score			**80.6%**

Assumed normalized scored for this example in the other scenarios
- Normalized score in CVNA-25 scenario: 76.7%
- Normalized score in CVNA-75 scenario: 100.0%
- Normalized score in CVNC scenario: 45.3%

Averaged AEB score for the four scenarios = 75.7%

HMI score:
- De-activation of the system not possible with a single push on a button **2 points**
- No FCW at speeds over 40 km/h **0 points**
- System switches off at low ambient lighting conditions **0 points**

Based on the above, the normalized HMI score = 50.0%

In total, the AEB Pedestrian total score will be:

$$AEB\ Pedestrian\ total\ score = (AEB\ score\ x\ 5) + (HMI\ score\ x\ 1)$$
$$= 5.0 \text{ x } 75.7\% + 1.0 \text{ x } 50.0\%$$
$$= 4.285 \text{ points}$$

Finally, the AEB Pedestrian score is included in the overall rating for the vehicle within the Pedestrian Protection box. It should be noted that Euro NCAP will only include the AEB Pedestrian score when the total passive safety protection score (headform, upper legform and lower legform) is 22 points or higher.

DISCUSSION
With the introduction of a relatively simple test to assess advanced systems like AEB, Euro NCAP would like to promote the introduction of these systems in the market. From the start of the development of the protocols, it was clear that the protocol would have to be reviewed and updated within a couple of years. For AEB VRU, additional cyclist scenarios will be included in 2018. On top of that, Euro NCAP will consider how and if it is

needed to also include other scenarios in daytime or obscure lighting conditions to the assessment of these systems.
The final decision whether the articulated pedestrian targets are to be used from 2016 onwards is yet to be made. As sensor systems get more advanced, the articulated dummy seems to support a more robust decision making process which improves system performance and acceptance.

All in all, Euro NCAP will continue to develop the requirements for AEB technologies to keep up with the development of these technologies and to ensure high quality systems for consumers.

CONCLUSIONS

Almost half of European road deaths in urban areas are pedestrians and cyclists. Vehicle design and technology such as forgiving front-ends and avoidance systems can help address this problem. In recent years, a number of vehicle manufacturers have started to offer AEB Pedestrian systems that mitigate the consequence of a potential crash with pedestrians and/or cyclist. To promote and guide the further development of these systems, Euro NCAP will adopt AEB Pedestrian systems in the rating from 2016 onwards. Based on the expected performance of current systems and accident priorities, test scenarios and an assessment scheme have been agreed between the main stakeholders in Europe. The proposed procedures are considered a first step and will be updated and expanded upon in the coming years.

ACKNOWLEDGEMENT

The P-NCAP working group was able to deliver the test and assessment protocols in time for implementation in 2016 due to all the valuable work done within and outside of the working group.

P-NCAP WG members

ADAC, BASt, IDIADA, NL-MOT/RDW, Swedish Transport Administration (STA), Thatcham, TNO and UTAC.

The members of the P-NCAP WG have all put significant effort and resources into the development and verification of these protocols. OEMs and suppliers are acknowledged for their support and feedback on the protocols, and special thanks goes out to the OEMs represented in ACEA/JAMA/KAMA for supporting the development of the protocols by funding and supporting round robin test series.

BEYOND SAFETY LEGISLATION: CONTRIBUTION OF CONSUMER INFORMATION PROGRAMMES TO ENHANCED INJURY MITIGATION OF PEDESTRIANS DURING ACCIDENTS WITH MOTOR VEHICLES

Oliver Zander
Federal Highway Research Institute (BASt)
Germany

Dirk-Uwe Gehring
BGS Boehme & Gehring GmbH
Germany

Michiel van Ratingen
European New Car Assessment Programme (Euro NCAP)
Belgium

Paper Number 15-0258

ABSTRACT

During the past five years, a Euro NCAP technical working group on pedestrian safety has been working on improving test and assessment procedures for enhanced passive pedestrian safety.

After harmonizing the tools and procedures as much as possible with legislation, the work was mainly focused on the development of grid procedures for the pedestrian body regions head, upper leg with pelvis and lower leg with knee. Furthermore, the test parameters for the head and the upper leg were revised, a new lower legform impactor was introduced and the injury thresholds were adjusted or, where necessary, the injury criteria were changed. Finally, the assessment limits and colour scheme were refined, widening the range and adding two more colours in order to provide a more detailed description of the pedestrian safety performance.

By abstaining from an assessment based on a worst point selection philosophy, the improved test point determination procedures that were introduced during the years 2013 and 2014 give a more homogeneous, high resolution picture of the pedestrian safety performance of the vehicle frontends. By using a uniform grid for each test zone approximately 200 test points, evenly distributed within each area, can now be assessed per vehicle.
The introduction of the flexible pedestrian legform impactor in 2014 enables a more realistic injury prediction of the knee and the tibia using a biofidelic test tool.

With the new upper legform test that has been launched in 2015 the assessment in that area is now focusing on the injured body region instead of the injury causing vehicle part and thus is aligned with the approach in the remaining body regions head and lower leg. At the same time, a monitoring test with the headform impactor against the bonnet leading edge is closing the possible gap between the test areas to identify injury causing vehicle parts that moved out of focus due to the introduction of the new upper legform test.

The paper describes the new test and assessment procedures with their underlying philosophy and gives an outlook in terms of open issues, specifying the needs for further improvement in the future.

In parallel to the work of the pedestrian subgroup, a Euro NCAP working group on heavy vehicles introduced a set of protocol changes in 2011 that were related to the assessment of M1 vehicles derived from commercial vehicles, with a gross vehicle weight between 2.5 and 3.5 tons and 8 or 9 seats. The paper also investigates the applicability of the new pedestrian test and assessment procedures to heavy vehicles.

INTRODUCTION

A Euro NCAP technical working group on pedestrian safety has elaborated during the past five years updated protocols for testing and assessment of passive pedestrian protection systems of passenger cars. After harmonizing the test tools and procedures as much as possible with pedestrian legislation, the overall objective of the group was to make the assessment of the pedestrian protection offered by the vehicle front-end more robust, reproducible, accurate and realistic. The work was mainly focused on the improvement of test procedures for the body regions head, pelvis, upper leg and lower leg, also including a new impactor for the lower leg test as well as a vehicle grid markup, the adaptation of injury criteria and a modified visualization by introducing a five colour scheme for all rated test procedures.

In parallel to the activities undertaken by the pedestrian working group, an ad hoc group of Euro NCAP developed specific test procedures for active systems of passive pedestrian safety, in particular active bonnets, that were as from version 5.3 onwards included within the pedestrian testing protocol. Another working group dealt with issues related to heavy M1 vehicles derived from commercial vehicles and the applicability of the new pedestrian test and assessment procedures to this class of vehicles.

This paper describes the main results of work undertaken by the pedestrian safety working group of Euro NCAP with input from the ad hoc working group on active pedestrian protection systems and the working group on heavy commercial vehicles that were successively implemented within the Euro NCAP test and assessment protocols for pedestrian protection.

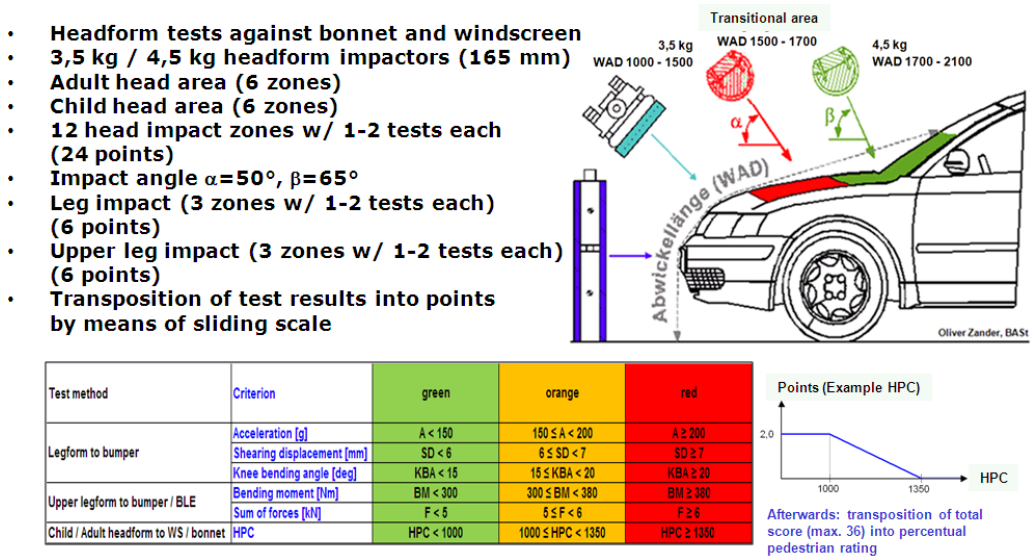
EURO NCAP UNTIL 2012

The first phase of work performed by the pedestrian safety group was mainly focused on a harmonization with Comission Regulation (EC) No. 631/2009 prescribing the test and assessment procedures for M1 vehicle type approval in Europe (European Commission, 2009). The borderline of test areas for child and adult headform impactors that were previously, in principle, described by wrap around distances (WAD) between 1000 and 1500 for the child headform impactor and between 1500 and 2100 for the adult headform impactor, was changed from WAD 1500 to WAD 1700, following the technical prescriptions of the Commission Regulation. Different to a hard borderline between both headfrom impact zones in legislation, Euro NCAP defined a transitional area between WAD 1500 and WAD 1700 where to use the adult headform impactor in the windscreen and windscreen base area and the child headform impactor on the bonnet and its periphery. In 2010, by introducing testing protocol version 5.0, the headforms themselves were changed from a 4.8 kg adult headform impactor with 165 mm diameter to a 4.5 kg impactor with unchanged dimensions, and from a 2.8 kg child headform impactor with a diameter of 130 mm to a 3.5 kg impactor with a diameter of 165 mm.

The philosophy of a worst point selection remained unchanged within Euro NCAP until the end of 2012: Both, the adult and the child headform areas were divided into six sixths, each of them subdivided into four quarters. While Euro NCAP selected the potentially most injurious impact point within each of the twelve sixths, the vehicle manufacturer was allowed to nominate up to three quarters within each sixth where then Euro NCAP again picked the expected hardest impact point. Altogether between 12 and 24 head impact points were selected, each sixth scored with a maximum of 2 assessment points. Although being assessed entirely, not all of these points were actually tested, given that points located on certain structures were defaulted either red (no score, e.g. on the A-Pillar) or green (full score, e.g. on the centre of the windscreen), symmetry occurred or that previous tests on adjacent areas already indicated the performance in the area to be assessed. For the upper and lower legform area, the method of worst point selection remained unchanged as well: both areas were divided into three thirds, each of them subdivided into two halves. Again, Euro NCAP selected the most injurious point within each third, while the manufacturer was given the possibility to nominate the remaining half of each third, where afterwards Euro NCAP again selected the hardest point. Altogether, between 3 and 6 impact points were determined in both, the upper and the lower legform area, each third scored with a maximum of 2 assessment points. Like in the headform area not all of these points were tested, given that symmetry occurred or there was agreement between the manufacturer and Euro NCAP of particular points defaulted red (no score) without being tested.

The last version of the Euro NCAP test and assessment protocols applying the philosophy of worst impact point selection is still valid nowadays as fall back scenario when the manufacturer opposes to the grid procedure that will be described in the following. In that case, the five colour scheme has to be applied.

An overview of the Euro NCAP test and assessment procedure before 2013 is depicted in Figure 1:



Test method	Criterion	green	orange	red
Legform to bumper	Acceleration [g]	A < 150	150 ≤ A < 200	A ≥ 200
	Shearing displacement [mm]	SD < 6	6 ≤ SD < 7	SD ≥ 7
	Knee bending angle [deg]	KBA < 15	15 ≤ KBA < 20	KBA ≥ 20
Upper legform to bumper / BLE	Bending moment [Nm]	BM < 300	300 ≤ BM < 380	BM ≥ 380
	Sum of forces [kN]	F < 5	5 ≤ F < 6	F ≥ 6
Child / Adult headform to WS / bonnet	HPC	HPC < 1000	1000 ≤ HPC < 1350	HPC ≥ 1350

Figure 1. Euro NCAP test and assessment procedure before 2013.

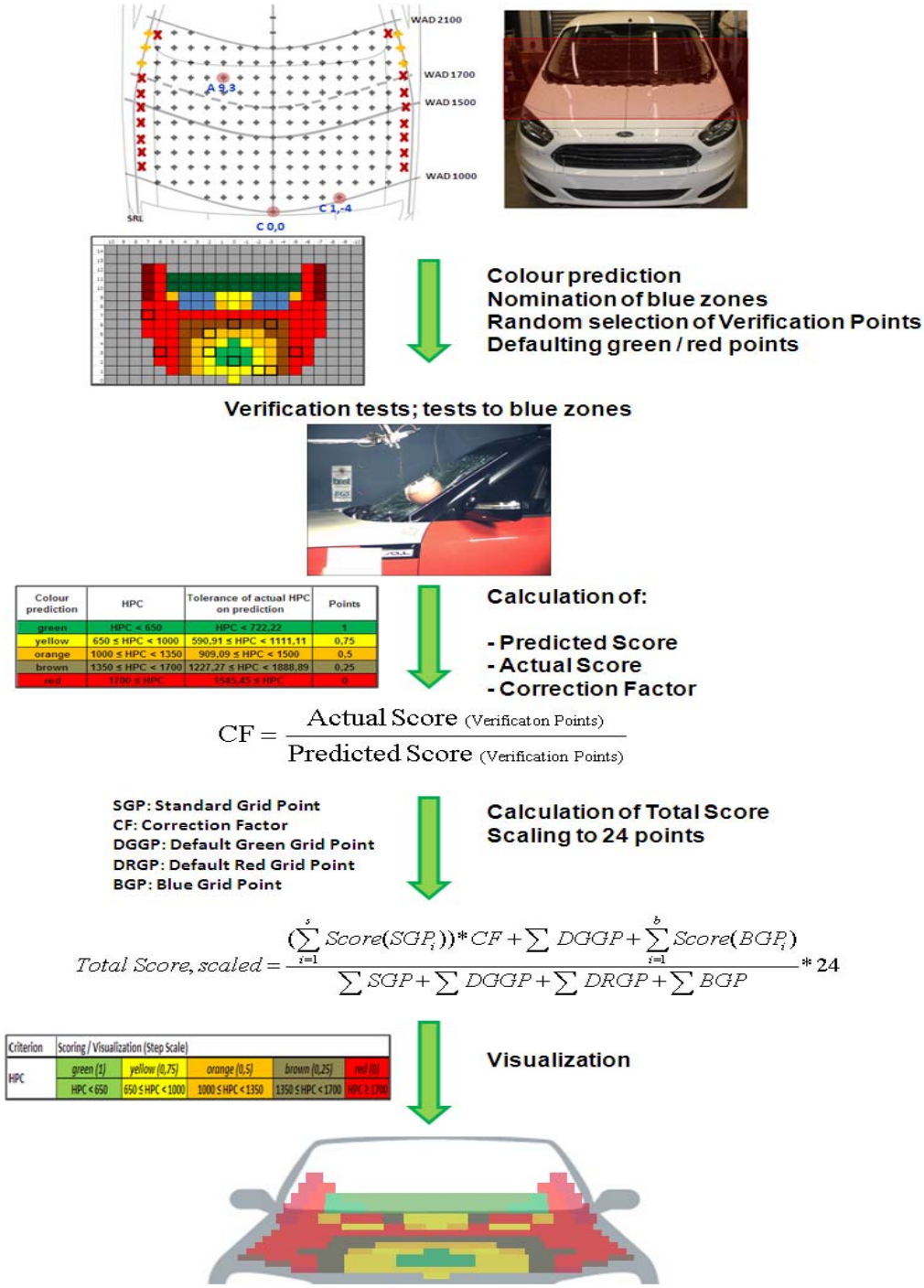
EURO NCAP 2013

The second phase of work performed by the pedestrian safety working group mainly concentrated on the grid procedure for the headform impactor tests. The principal idea of the new method was to abstain from a, to some extent, subjectively performed selection of worst case head impact points, towards an objective and homogeneous assessment of the entire headform area. By introducing the possibility of providing in-house simulation results by the vehicle manufacturer, the amount of tests was limited, keeping the testing effort in the previous range. At the same time, the sliding scale assessment procedure between HIC 1000 and HIC 1350 with three colours was changed to an incremental approach between HIC 650 and HIC 1700 and five color bands, the middle three of them having a range of HIC 350: HIC below 650 - green, 2 points; HIC between 650 and 1000 - yellow, 1.5 points; HIC between 1000 and 1350 - orange, 1 point; HIC between 1350 and 1700 - brown, 0.5 points; HIC over 1700 - red, no point.

According to the new method and starting from the intersection between the vehicle's vertical longitudinal centerplane and the WAD 1000 (C 0,0, see Figure 2) a grid with 100 mm * 100 mm resolution on the xy plane is vertically projected onto the headform area that is, in principle, described by the WAD 1000 and 2100 and the side reference lines, where all grid points within a distance of less than 50 mm to the side reference lines are deleted, and additional grid points are marked on the side reference lines on the A-pillar. After the provision of the safety performance by means of colour information for each of the determined grid points by the OEM, a default minimum of 10 and, on request of the vehicle manufacturer, a maximum of 20 verification points is generated randomly. Additionally, the manufacturer is given the opportunity to select a maximum of 8 blue zones, each consisting of 1 or 2 grid points, where the safety performance is unknown or previous tests have indicated instable results, and which are tested once on the point selected by Euro NCAP, unless symmetries are being applied. Thus, in total a minimum of 10 and a maximum of 28 head impact tests are to be performed according to the headform grid method. After testing, a correction factor is calculated by the quotient of the sum of actual verification test results and the sum of the points resulting from the colour predictions. The correction factor provides an indication of how accurate actual testing matches the prediction and should be between 0.75 and 1.25 for the predictions to be accepted by Euro NCAP. By multiplying the sum of points obtained by the colour predictions times the correction factor, the actual performance of the predicted grid points is calculated.

In a last step, the total head score is calculated by including the number of green defaulted points and the actual test result from the blue zones, and scaled to the 24 points that are available for the head performance in box 3 (pedestrian protection) of the Euro NCAP overall rating scheme. A visualization of the test results by applying a five colour code scheme defined by the underlying results to each individual grid point completes the head assessment.

A flowchart describing the principal markup, test and assessment procedures for the headform grid is illustrated in Figure 2:



Colour prediction	HPC	Tolerance of actual HPC on prediction	Points
green	HPC < 650	HPC < 722,22	1
yellow	650 ≤ HPC < 1000	590,91 ≤ HPC < 1111,11	0,75
orange	1000 ≤ HPC < 1350	909,09 ≤ HPC < 1500	0,5
brown	1350 ≤ HPC < 1700	1227,27 ≤ HPC < 1888,89	0,25
red	1700 ≤ HPC	1545,45 ≤ HPC	0

Criterion	Scoring / Visualization (Step Scale)				
HPC	green (1)	yellow (0,75)	orange (0,5)	brown (0,25)	red (0)
	HPC < 650	650 ≤ HPC < 1000	1000 ≤ HPC < 1350	1350 ≤ HPC < 1700	HPC ≥ 1700

Figure 2. Euro NCAP headform grid procedure.

EURO NCAP 2014

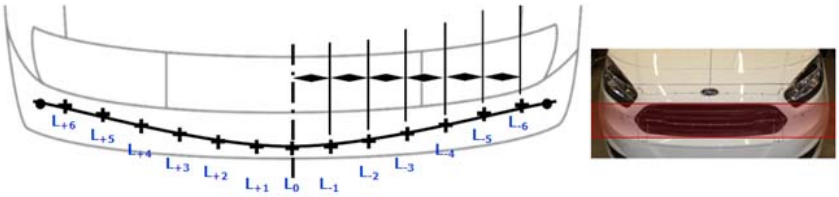
The next big step in the further development of the Euro NCAP pedestrian test and assessment procedures was finalized with the introduction of the flexible pedestrian legform impactor (FlexPLI) from the beginning of the year 2014 on. The FlexPLI with enhanced biomechanical properties compared to those of the lower legform impactor according to EEVC, in particular in the knee and tibia area, was developed by the Japan Automobile Research Institute (JARI) as from the year 2000 onwards and evaluated by a technical evaluation group (TEG) under the umbrella of the Working Party on Passive Safety (GRSP) of the United Nations Economic Commission for Europe (UNECE) since 2005. Last issues with the impactor were dealt with by an Informal Working Group on Phase 2 of GTR9 before UNECE Working Party 29 finally adopted the FlexPLI, in a first step, for the 01 series of amendments of UN Regulation No. 127 on Pedestrian Safety (UNECE, 2015) that has become effective on 22 January 2015. Until the mandatory use of the FlexPLI as from 1 September 2017 onwards, the vehicle manufacturers are given the choice to either submit new vehicles to type approval using the new test tool or to alternatively use the lower legform impactor developed by Working Group 17 of the European Enhanced Vehicle-safety Committee (EEVC).

With the implementation of tests with the FlexPLI, the Euro NCAP markup and test procedures changed from the subdivision of the upper and lower legform test area into six subareas each to a grid markup, testing every second grid point and allocating the test result to the corresponding adjacent grid points; besides, where possible, application of symmetry.

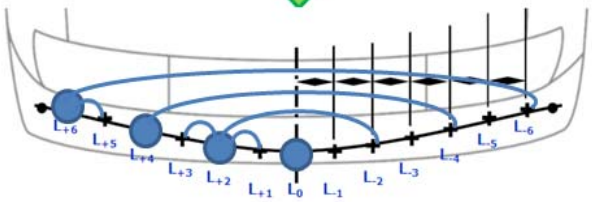
For the lower legform tests with the FlexPLI, starting at y_0 (vehicle centreline), a grid with 100 mm resolution is marked in lateral direction onto the upper bumper reference line. If the distance between the last (outermost) grid point and the end of the test area is greater than 50 mm, an additional grid point is marked at a lateral distance of 50 mm to the last grid point onto the upper bumper reference line. After defining the starting point which is located either on y_0 or one of its adjacent points, every second grid point is selected and tested. Where possible, symmetries are to be applied. All asymmetrical grid points not being tested are awarded with the worse of the two results coming from both adjacent grid points. Prior to the determination of the starting point by Euro NCAP, the vehicle manufacturer is given the choice to nominate grid points to be exempted from being assessed by taking over the results from adjacent or symmetrical identical grid points, thus these grid points need to be tested.

Along with the introduction of the FlexPLI in Euro NCAP, Zander (2011) derived upper performance limits from the injury criteria of the FlexPLI as equivalents for 20% tibia fracture risk, 15 deg knee bending angle of the EEVC WG 17 pedestrian legform impactor and a transposition of the results of PMHS testing reported by Bhalla et al. (2003) to the FlexPLI. In terms of assessment of the lower legform test results, the pedestrian subgroup decided to equally balance the injury risks related to the medial collateral ligament (MCL) and the tibia segments, whereas points for the MCL are only awarded in case of not exceeding the identified risk of 10 mm elongation for cruciate ligament rupture (ACL/PCL). Furthermore, as far as the tibia is concerned, only the highest of the four bending moments is taken into account for the assessment. Between the defined upper and lower performance limits for MCL elongation and tibia bending moment, a sliding scale is applied to both criteria. The total score for the lower legform area is calculated by adding the points for the individual grid points and scaling the results to 6 points. For visualization, a five colour scheme is applied.

Figure 3 describes in principle the markup, test and assessment procedure for the tests with the FlexPLI:



OEM to nominate points to be tested in either case
Definition of starting point (L_0 or L_{+1})
Use of symmetry, where applicable
Testing of every second grid point
Allocation of results to adjacent points



Criterion	Full score	Sliding Scale	0	Points
Tibia BM [Nm]	$BM_{max} < 282$	$282 \leq BM_{max} < 340$	$BM_{max} \geq 340$	0 - 0,5
MCL-EL [mm]	$EL_{MCL} < 19$	$19 \leq EL_{MCL} < 22$	$EL_{MCL} \geq 22$	0 - 0,5
ACL/PCL-EL [mm]	$EL_{ACL/PCL} < 10$		$EL_{ACL/PCL} \geq 10$	n/a

Calculation of grid point score:

- Tibia BM & MCL (use 10 mm ACL/PCL switch)
- Use of sliding scale
- Weighting of tibia and knee (50%/50%)

$$Score(GP) = Score(TibiaBM_{max}) + Score(MCL)$$

Calculation of total score
Scaling to 6 points

$$\frac{\sum_{i=1}^{n} Score(GP_i)}{n} * 6$$

Score (GP)	S=1	1>S>0,75	0,75≥S>0,5	0,5≥S>0,25	0,25≥S

Visualization

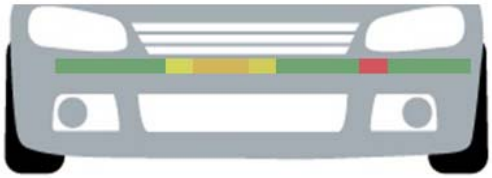


Figure 3. Euro NCAP FlexPLI grid procedure.

In 2014, the test procedure and the injury criteria for the upper legform impactor remained in principle unchanged. For the markup, however, a grid along with a five colour scheme based on grid point scores as for the headform and the lower legform tests was inroduced. Besides, in 2014 only, a requirement was applied that in case of the distance between the last (outermost) grid point and the corner reference point being greater than 75 mm, an additional grid point was to be marked at a lateral distance of 50 mm to the corner reference point onto the bonnet leading edge reference line.

EURO NCAP 2015

In 2015, Euro NCAP introduced two new pedestrian protection tests, namely the new upper legform test replacing the upper legform to bonnet leading edge test, and the headform to bonnet leading edge test.

Upper legform test

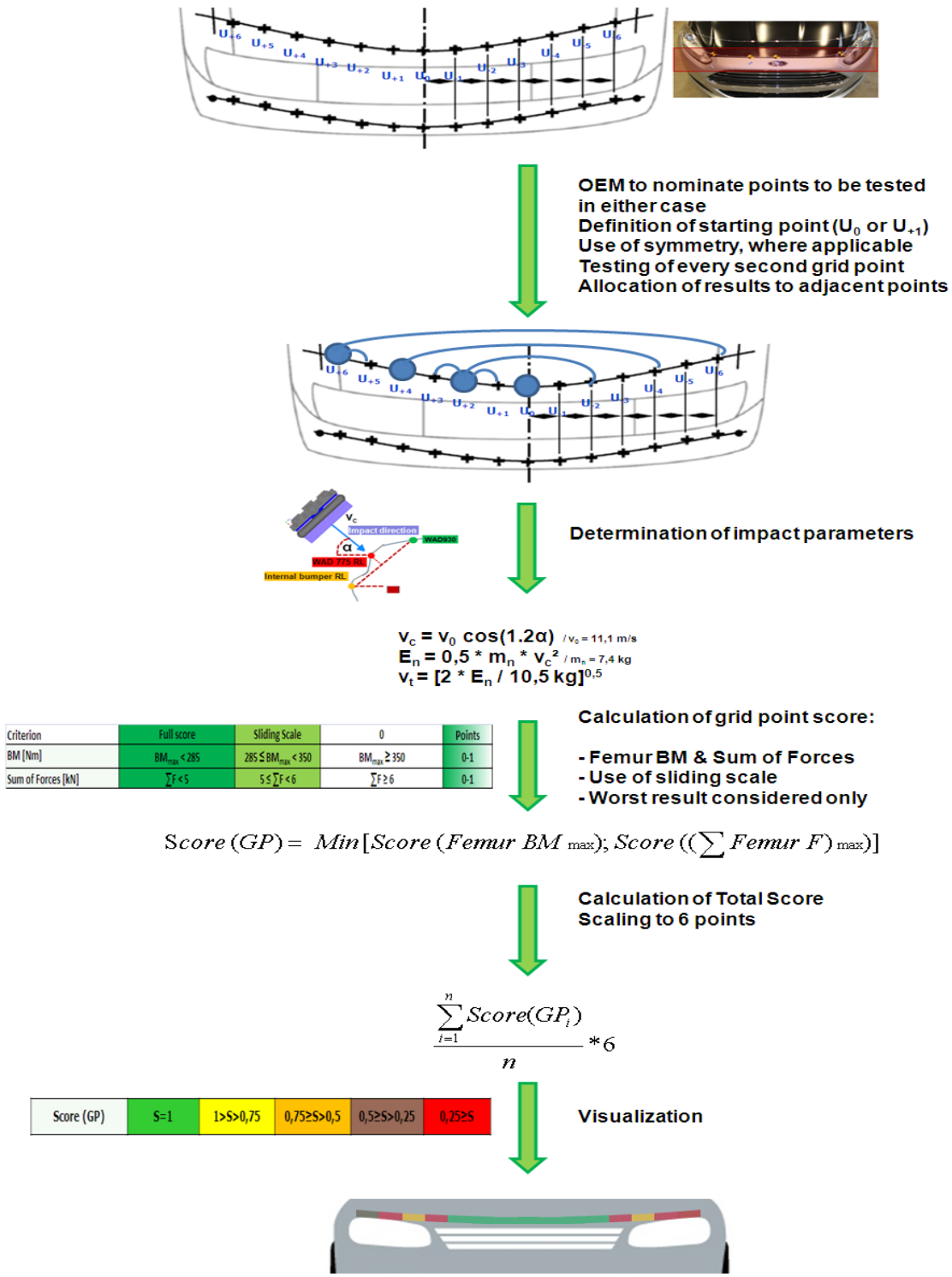
As from 2015 onwards, the test with the upper legform impactor was fundamentally changed. Main reason for the significant change was an analysis of accident data showing a decreasing number of upper leg injuries caused by the bonnet leading edge, mainly due to the development of passenger vehicle frontends away from the typical Sedan design with sharp and high leading edges towards rounded frontends with lower and significantly softened leading edges. Besides, a change of injury patterns could be observed. Thus, a test with the upper legform impactor against the bonnet leading edge was no longer seen as best representing present real world injury data. However, real world accident data also showed that injuries of femur and pelvis are still of a high importance and therefore need to be addressed. Therefore, since this point in time, the injured body regions (pelvis, femur) rather than the injury causing vehicle parts are in the focus of assessment. Thus, the new test procedure is in line with the assessment of head and lower leg injuries. The new upper legform test aims for a more realistic simulation of the correct impact height and position of the human thigh and pelvis during an impact. Therefore, the nominal impactor weight is standardized to 7.4 kg and the upper load cell of the impactor is approximately aligned with WAD 930 as the height of the human hip joint. As the impactor centerline always aims at WAD 775, the corresponding WAD line is used as markup line for the upper legform grid points. The angle of impact α depends on the outer vehicle contour and is perpendicular to a line on the same vehicle vertical longitudinal plane as the respective grid point, connecting the WAD 930 with the internal bumper reference line (IBRL). The IBRL is a connection line of the averaged grid point bumper beam heights (i.e. averaged height of the bumper beam at the lateral grid point position and its two adjacent markings at a distance of 33,3 mm from the grid point), marked on the bumper beam and projected onto the bumper fascia. The nominal vehicle speed v_0 at time of first pedestrian contact, the relative speed v_c and the angle of impact α of the upper leg, both measured at time of mid femur contact with the vehicle front, are used for calculation of the test speed v_t under application of a uniform impactor test mass of 10.5 kg (Figure 4). These calculations lead to possible impact angles of up to 44.7 degrees and maximum test speeds of 9.3m/s. As a comparison with the previous energy calculation procedure shows that the impact energy is reduced for many vehicles, the lower energy limit is reduced to 160 J, i.e. tests will now also be performed with energies below 200 J up to a minimum of 160 J. The upper energy limit of 700 J is not needed anymore because the results from the energy calculation are limited to 456 J.

For the upper legform test, starting at y_0, a grid with 100 mm resolution is marked in lateral direction onto the WAD 775 reference line up to a distance of no less than 50 mm to the lateral projection of the corner reference point. After defining the starting point which is located either on y_0 or one of its adjacent points, every second grid point is selected and tested. Again, whereever possible, symmetries are to be applied. All asymmetrical grid points not being tested are awarded with the worse of the two results coming from both adjacent grid points. As for the lower legform tests, prior to the determination of the starting point by Euro NCAP, the vehicle manufacturer is given the opportunity to nominate grid points to be exempted from being assessed by taking over the results from adjacent or symmetrical identical grid points which thus must be tested.

Regarding the test evaluation, in contrary to the tests with the FlexPLI, no weighting of the maximum bending moments and the maximum sum of forces is done, but, as before, the worst of the two results is taken into account only. For the allocation of points to the particular test results, a sliding scale is applied between the upper and lower performance limits, which in terms of the maximum bending moments has been decreased to 285 or 350 Nm respectively.

The total score for the upper legform area is calculated by adding the points for the individual grid points and scaling the results to 6 points. For visualization, the five colour scheme that was already introduced in 2014 remains unchanged.

The flowchart of the upper legform markup, test and assessment procedure is shown in Figure 4:



Criterion	Full score	Sliding Scale	0	Points
BM [Nm]	$BM_{max} < 285$	$285 \leq BM_{max} < 350$	$BM_{max} \geq 350$	0-1
Sum of Forces [kN]	$\sum F < 5$	$5 \leq \sum F < 6$	$\sum F \geq 6$	0-1

Score (GP)	S=1	1>S>0,75	0,75≥S>0,5	0,5≥S>0,25	0,25≥S

Figure 4. Euro NCAP New Upper Legform grid procedure.

Headform to bonnet leading edge test

With the introduction of the new upper legform test, the bonnet leading edge is not anymore in the focus of Euro NCAP passive pedestrian safety assessment. Longhitano et al. (2005) and Zander (2014) found the bonnet leading edge having a high relevance as injury causing part in real world accident data. While in most cases, the bonnet leading edge is to some extent also covered by the new upper legform test, in case of being located between WAD 930 and WAD 1000 the bonnet leading edge reference line (BLE-RL) is not assessed anymore, see Figure 5.

Figure 5. Untested critical area after introduction of Euro NCAP New Upper Legform grid procedure.

To encounter the injury risk for vulnerable road users, in particular head injuries of small children, that is related to the bonnet leading edge, BASt had proposed to the Euro NCAP pedestrian safety working group a supplementary test with the child headform impactor against the bonnet leading edge as illustrated in Figure 6. The proposal was adopted by Euro NCAP and is detailed in Technical Bulletin 019 (Euro NCAP, 2014). For the vertical longitudinal plane of any upper legform grid point where the bonnet leading edge reference line is located between WAD 930 mm and WAD 1000 mm, an additional test with the child headform impactor is performed at the intersection of the vertical longitudinal plane with the bonnet leading edge reference line at an impact speed of 40 km/h under an impact angle of 20° ±2° to the Ground Reference Level. The result of this test will be monitored against a HIC value of 650.

Figure 6. Headform to bonnet leading edge test.

ACTIVE PEDESTRIAN PROTECTION SYSTEMS OF PASSIVE PEDESTRIAN SAFETY

An ad hoc Euro NCAP working group on active pedestrian protection systems has developed test and assessment procedures for deployable (non-static) systems, in particular pop up bonnets. As the legal requirements in that area do not cover significant aspects related to a pedestrian impact on a deployable system, the working group defined additional test and assessment procedures such as requirements for sensor systems and actuators to detect the hardest to detect pedestrian, requirements in terms of the total response time (TRT) of the system (time between the first contact of the pedestrian and full deployment of the safety system), performance requirements below the deployment threshold as well as at higher speeds after initiation of the deployment. Furthermore, the bonnet deflection due to pedestrian torso contact at pedestrian head impact location at time of the impact is required to not exceed a certain extent. Finally, the grade of fulfillment of the requirements defines the ambient conditions and test parameters, i.e. whether a system is tested in closed position, in fully deployed position or dynamically. The active pedestrian protection test procedures have been implemented within the test protocol version 5.3 and are valid since January 2011.

HEAVY VEHICLES

In 2011, the Euro NCAP working group on heavy vehicles finalized the development of a set of changes to the different protocols that were particularly related to the testing and assessment of M1 vehicles derived from commercial vehicles, with a gross vehicle weight between 2,5 and 3,5 tons and 8 or 9 seats. In terms of pedestrian safety, BASt investigated the main differences between passenger cars and heavy commercial vehicles and proposed to the working group a set of particular test parameters that were different to the ones related to the standard pedestrian tests (Zander 2010-1 and 2010-2). Essentially, as main differences, Euro NCAP concluded a uniform head impact angle of 50 degrees rearwards and of 20 degrees forward of the bonnet leading edge reference line. Besides, at impact points with a height of the bonnet leading ede reference line greater than 835 mm, no test with the upper legform impactor against the BLE was required. In terms of rating, a soft landing for heavy vehicles was introduced: while in 2011, a fufilment rate of 25 percent was required for 5 stars balancing criteria, the requirements were identical between commercial vehicles and passenger cars for four and five stars vehicles in 2014.

After finalization of the pedestrian test procedures for 2015, the Euro NCAP working group on pedestrian safety discussed the applicability of the modified tests to commercial vehices. Gehring et al. (2014) examined several heavy commercial vehicles in detail (see Figure 7) and found the general applicability of the pedestrian testing protocol also to heavy vehicles.

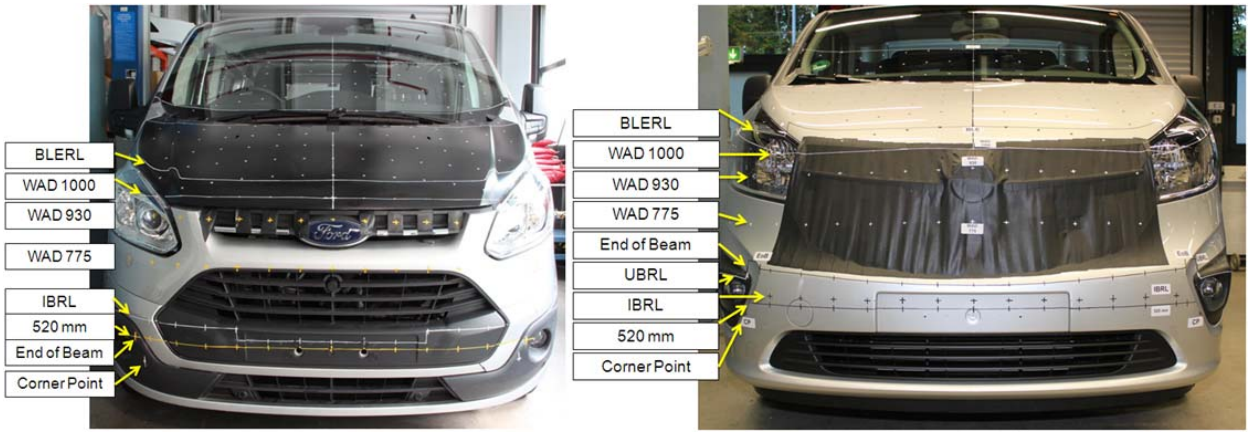


Figure 7. Investigation on the applicability of pedestrian testing protocol to heavy vehicles.

As in the meantime a head impact angle of 20 degrees for tests on or forward of the BLERL had already been introduced for passenger cars, the remaining main difference to the pedestrian testing protocol was the adult headform angle of 50 degrees instead of 65 degrees for the passenger cars. During the investigations, Gehring et al. also found some potential for improvement that was well applicable to the pedestrian testing protocol. For practical reasons they proposed for steeper parts of the outer contour of the vehicle to markup the grid using a horizontal instead of a vertical projection. This proposal was adopted by Euro NCAP for all areas forward of the BLERL where the angle of a line, connecting the BLE-RL and WAD 1000 on the vehicle's vertical longitudinal centerplane, is greater than 60 degrees to the ground reference level. For relevant gaps in the markup area, the outer contour of the vehicle is to be approximated with tape, whereas between the lower bumper reference line and the BLE, a wrap around is to be created.

Due to the fact that within the new upper legform test the impactor centreline is always aligned with WAD 775 and thus the grid point height is always below 835 mm, the test exemption that was valid for the first phase of the heavy vehicles protocol is not applicable anymore with the introduction of the new test.

DISCUSSION

In the past five years, the work performed by the Euro NCAP pedestrian safety working group with input from the ad hoc working group on active pedestrian protection systems as well as the heavy vehicles working group was mainly focused on three topics related to pedestrian safety. The first topic concentrated on harmonization with the legal requirements. In this context, amongst other things, the headform impactors and impact areas as well as the lower performance limits were aligned with the impactors and with the pass/fail criterion for one third of the impact area detailed in Commisson Regulation (EC) No. 631/2009.

The second topic dealt with the objectification of the test procedures which was mainly done by introducing a high resolution grid over the entire front area to be assessed, resulting in detailed and homogeneous information in terms of the pedestrian safety performance. The grid approach has opened up the protocol to virtual testing predictions by vehicle manufacturers based on CAE. This way of combining physical testing and simulations is a first in consumer safety testing. The first two years of experience with the head grid method give evidence that the capability of predicting the safety performance is already at a high level. A study of the results of 70 vehicles tested and assessed according to the 2013 and 2014 protocols shows that the target of a correction factor between 0,75 and 1,25 was met in all but three cases. In most cases, the correction factor was significantly below the phase-in target of a deviation from the actual verification test results of no more than 25 percent. With the majority of cars having a positive correction factor above 1, the mean value was 1,05, meaning only a 5 percent underestimation by the given colour predictions, being well in line with the actual safety performance, proving the feasibility and applicability of the grid method using colour predictions.
Also testing active systems of passive pedestrian safety such as active bonnets is allowing human body model simulations as part of the evidence, what again is a new approach in the world of safety assessment and consumer testing.

The third topic was related to improved biomechanics of the test tools and procedures. With the introduction of the FlexPLI a biofidelic test tool with the capability of simulating the kinematic behaviour and human responses of the lower leg and knee area in an appropriate manner along with new injury criteria was introduced. Here, the first year of testing resulted in a majority of cars performing quite well in the lower leg area. A comparison of the 2013 and 2014 lower legform results gives evidence of the feasibility of the new test method using the FlexPLI. While in 2013 the 31 assessed vehicles resulted in an average of 5.7 points when tested with the EEVC WG 17 lower legform impactor, the mean value of the 39 vehicles tested in 2014 with FlexPLI and grid method was 5,9 points. In 2013, 87% of the vehicles scored full points in the lower legform area; in 2014, 76 % of vehicles had entire green bumper zones. None of the assessed vehicles scored less than 4.8 points.
The upper legform impactor remained, in principle, without further modification but with the focus on the injured body regions pelvis and femur rather than on the bonnet leading edge as injury causing part of the vehicle. To avoid crucial gaps in the assessment, the bonnet leading edge is covered by an additional test with the headform impactor, where necessary.

A fourth topic that has not yet been entirely covered by the work in the past years is the review of pedestrian injury patterns currently existing in real world accidents. Figure 8 depicts injury patterns occurring in vehicle to pedestrian collisions according to AIS98 code based on body parts (Zander et al., 2015):

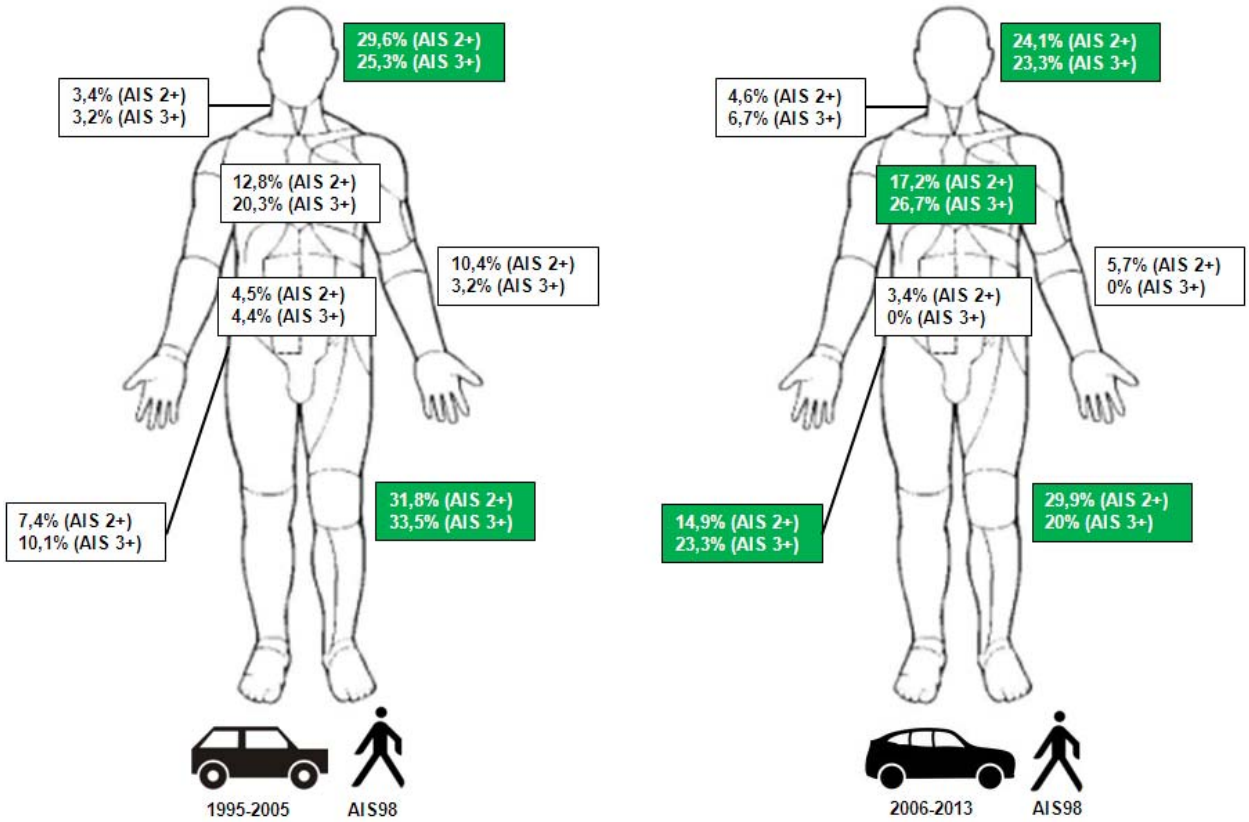


Figure 8. Injury patterns occurring in vehicle to pedestrian collisions according to AIS98 code based on body parts. The Figure shows the relations between the injury severity from AIS 2 (or AIS 3 respectively) to AIS6 (i.e. AIS2+ or AIS3+ respectively) in the particular body region and AIS2+ (AIS3+) injury severities of all body regions.

The in-depth investigation of road accidents in Germany shows the relevance of AIS2+ and AIS3+ head and leg injuries (approximately 30% each) in particular for pedestrians involved in accidents with passenger cars registered between 1995 and 2005, i.e. in the time when the pedestrian impactors have been developed by the pedestrian safety working groups of the European Enhanced Vehicle-safety Committee (EEVC), but still before application of the Framework Directive 2003/102/EC as predecessor of Commission Regulation (EC) No. 631/2009. During accidents with passenger cars registered from 2006 onwards, i.e. after application of the Framework Directive, the distribution especially of the AIS3+ coded injuries has been shifted towards the body regions thorax and pelvis while the relevance for the head and leg remains more or less unchanged. From this study it can be concluded that on the one hand, a test procedure covering the pelvis area as now introduced by Euro NCAP is addressing the real world accident scenarios, but also that an additional thorax test could be expected to significantly reduce thoracic injuries on the other hand.

Another point that has not yet been addressed is the safety of cyclists being the second big group of vulnerable road users besides pedestrians. Euro NCAP is planning to cover cyclist safety as from 2018 onwards by means of active safety. However, like for pedestrians, a baseline passive safety level also for cyclists should be ensured, especially as active safety systems are expected to mitigate rather than avoid crashes in most cases. A current research project at BASt dealing with the development of passive test and assessment procedures for cyclists aims at implementing modifications to the current pedestrian procedures to adequately cover both groups of vulnerable road users.

Finally, in the light of demographical change the procedures to-be should also focus on elderly as vulnerable road users. The upcoming European SENIORS (Safety-ENhancing Innovations for Older Road userS) project settled under the European HORIZON 2020 framework programme will in particular address the passive safety of the elderly.

CONCLUSIONS

The Euro NCAP pedestrian test and assessment protocols have been substantially updated and, where necessary, amended. Where possible, the procedures were harmonized with legislation. The assessment has gained objectivity by introducing grid procedures for all assessed body regions. The tools and test procedures have been modified to better reflect impact kinematics as well as the biomechanical response that can be observed in real world accidents. However, there is room for further improvement in terms of real world injuries and their coverage by the latest test and assessment procedures. The linearly guided upper legform impactor is lacking biofidelic behavior that should be reflected in terms of pedestrians' pelvis and femur. While knee and tibia injuries can be reliably predicted with the FlexPLI, the readings for the femur cannot be used mainly due to the lack of the torso mass of the pedestrian. Studies by Zander et al. (2011) have shown a good correlation between full scale dummy tests and tests with the FlexPLI with applied upper body mass (UBM); however, for an implementation within the existing test procedures the generated database needs to be extended by more vehicle frontends and human body simulations, i.e. further research is needed. In depth accident studies show a redistribution of pedestrian injury patterns towards thoracic and pelvic injuries that need to be addressed by new or improved impactors and test procedures. Furthermore, the procedures need to also cover cyclists as the second big group of vulnerable road users. Finally, pedestrian safety should also take care of the elderly being the most vulnerable road users. Results in this context can be expected from the upcoming European SENIORS project.

REFERENCES

Bhalla K., Bose D., Madeley N., Kerrigan J., Crandall J. 2003. „Evaluation of the response of mechanical pedestrian knee joint impactors in bending and shear loading.“ Paper no. 429 of ESV conference proceedings.

European Union. 2009. "COMMISSION REGULATION (EC) No 631/2009 of 22 July 2009 laying down detailed rules for the implementation of Annex I to Regulation (EC) No 78/2009 of the European Parliament and of the Council on the type-approval of motor vehicles with regard to the protection of pedestrians and other vulnerable road users, amending Directive 2007/46/EC and repealing Directives 2003/102/EC and 2005/66/EC." Official Journal of the European Union, 25 July 2009.

European Union. 2003. "Directive 2003/102/EC of the European Parliament and of the Council of 17 November 2013 relating to the protection of pedestrians and other vulnerable road users before and in the event of a collision with a motor vehicle and amending Council Directive 70/156/EEC." Official Journal of the European Union, 06. December 2003.

European New Car Assessment Programme. 2011. "Pedestrian Testing Protocol Version 5.3.1." Euro NCAP, November 2011.

European New Car Assessment Programme (Zander O., van Ratingen M.). 2014. "Headform to Bonnet Leading Edge Tests. Technical Bulletin TB 019 Version 1.0." Euro NCAP, May 2014.

Gehring D.-U., Zander O. 2014. "Application of the new Euro NCAP pedestrian test procedures to heavy vehicles." Presentation to the Euro NCAP Pedestrian Safety Working Group during their 30th meeting. 25 November 2014.

Longhitano D., Ivarsson J., Henary B., Crandall J. 2005. "Torso injury trends for pedestrians struck by cars and LTV's." Paper no. 05-0411 of 19th ESV conference proceedings. Washington DC, 2005.

United Nations Economic Commission for Europe. 2015. “Agreement Concerning the Adoption of Uniform Technical Prescriptions for Wheeled Vehicles, Equipment and Parts which can be Fitted and/or be Used on Wheeled Vehicles and the Conditions for Reciprocal Recognition of Approvals Granted on the Basis of these Prescriptions (Revision 2, including the amendments which entered into force on 16 October 1995). Addendum 126: Regulation No. 127 01 series of amendments to the Regulation – Date of entry into force: 22 January 2015. Uniform provisions concerning the approval of motor vehicles with regard to their pedestrian safety performance.” Website of the UNECE, Doc. ECE/TRANS/505/Rev.2/Add.126/Rev.1. 4 February 2015

Zander O. 2010-1. “Testing of heavy vehicles according to a draft pedestrian testing protocol.” Presentation to the Euro NCAP Heavy Vehicles Subgroup during their 6th meeting. Brussels, 9 September 2010.

Zander O. 2010-2. “Proposal for BLE testing of heavy vehicles.” Presentation to the Euro NCAP Heavy Vehicles Subgroup during their 7th meeting. Brussels, 8 October 2010.

Zander O. 2011. “Introduction of FlexPLI within Euro NCAP legform test procedure.” Presentation to the Euro NCAP Pedestrian Protocol Subgroup during their 13th meeting. Brussels, 6 May 2011.

Zander O., Gehring D., Leßmann P. 2011. “Improved assessment methods of lower extremity injuries in vehicle-to-pedestrian accidents using impactor tests and full scale dummy tests.” Paper no. 11-0079 of 22nd ESV conference proceedings.

Zander O. 2014. “Status of the Lower Extremity Test and Assessment Procedures for Vulnerable Road Users. Proceedings of the 9th Praxis Conference Pedestrian Protection, Bergisch Gladbach, 2014.

Zander O., Wisch M. 2015. “Kausalzusammenhänge bei Radfahrer-Pkw-Unfällen, Verletzungsmuster von Radaufsassen und Maßnahmen von Gesetzgebung und Verbraucherschutz zur Minderung der Verletzungsschwere.“ Annual Report 2013-2014 of the Federal Highway Research Institute (BASt). 2015

EVALUATION OF THE EURO NCAP WHIPLASH PROTOCOL USING REAL-WORLD CRASH DATA

Anders Kullgren
Folksam Research and Chalmers University of Technology
Sweden

Brian Fildes
MUARC, Monash University
Australia

Michiel van Ratingen
James Ellway
Euro NCAP, Belgium

Michael Keall
University of Otago
New Zealand

Paper Number 15- 0267

ABSTRACT

Whiplash injuries account for the vast majority of casualties in road traffic crashes, leading to long-term consequences. The majority occur in rear-ends crashes. Consumer crash tests play an important role in promoting effective concepts to reduce the problem. The current Euro NCAP whiplash test protocol includes three sled tests at varying impact speeds and pulse shapes using a BioRID test dummy and 8 measures to assess whiplash potential based on previous best practice. Given the complexity of the test and with more experience, a real-world evaluation of the current protocol was undertaken. Three analyses were undertaken comprising an analysis of test outcome data, a logistic regression analysis, a ROC analysis, and a correlation analysis comparing crash and injury outcome. 13,389 drivers reporting whiplash injury symptoms to Folksam Insurance in Sweden were studied, of which 1,266 occurred in cars tested by Euro NCAP. For all occupants reporting initial symptoms, the risk of permanent medical impairment was followed up according to the procedure used by Swedish insurance companies. Test scores according to Euro NCAP, JNCAP and IIWPG protocols were calculated, as well as combinations of the three Euro NCAP pulses. For each combination or protocol, the test score was compared with the real-world outcome. A correlation analysis of the included injury criteria was also performed for the three crash pulses included.
The results showed that overall Euro NCAP, JNCAP and IIWPG all predict real-world whiplash injury outcome in terms of Permanent Medical Impairment (PMI). Based on limited data available, there was no statistical evidence using logistic regression and ROC analyses that any of the three tests performed better than any other. Correlations between the test scenarios of each of the three protocols, as well as the outcome associations with crash outcomes, suggested consistent improvements in the risk of permanent medical impairment. The main strength of the analyses conducted here was to show the validity of Euro NCAP, JNCAP and IIWPG whiplash test protocols when measured against real-world crash outcomes, which are the most important criteria showing that the tests are appropriately designed to help prevent injuries among the community. Some caution needs to be taken with these findings as many were not statistically significant because of the limited number of cases available. Further evaluation when additional data are available is warranted.

INTRODUCTION

Soft tissue injuries to the neck and associated areas, so called whiplash associated disorders (WAD), in rear-end crashes are a major societal burden in terms of reported incidence, on-going disability, cost, and inability to return to work (Krafft 1998, Malm et al. 2008, Kullgren et al. 2013). These injuries account for half of all injuries leading to long-term or permanent medical impairment (Krafft 1998, Kullgren et al. 2013). WAD often occur in relatively low severity crashes, typically at a change of velocity below 25 km/h (Eichberger et al. 1996, Kullgren et al. 2003), and in all impact directions, although WAD in rear impacts are most frequent (Krafft 1998, Watanabe et al. 2000, Kullgren et al. 2013).

Since the late 1990s more advanced concepts aimed at reducing the risk of WAD have been introduced on the market (Jakobsson 1998, Wiklund and Larsson 1997, Lundell et al. 1998, Sekizuka 1998). Better protection is achieved through improved geometry and dynamic properties of the head restraint or by active devices that move in a crash as the body loads the seat. The main ways to lower the risk of WAD in rear impacts are to minimize the relative motion between head and torso, to control energy transfer between the seat and the human body and to absorb energy in the seatback. Studies have been presented showing the effect of the some seat concepts indicating a reducing effect in WAD of approximately 20-50% (Viano and Olsén 2001, Farmer et al. 2008, Jakobsson 2004, Jakobsson et al. 2008, Krafft et al. 2003, Kullgren et al. 2013).

In 2004 consumer test programmes were introduced (IIWPG and Folksam/SRA) (Thatcham 2013, Krafft et al. 2004). And since 2009 the European New Car Assessment Program (Euro NCAP) has introduced a test protocol for assessing the Whiplash potential of vehicle seats as part of its vehicle safety rating system in Europe. The protocol specifies 3-sled tests at varying impacts using a BioRID test dummy and 8-measures to assess whiplash potential based on previous best practice. To date, more than 200 make and model vehicles seats have been evaluated using the Euro NCAP protocol. JNCAP first conducted similar assessments of vehicles in Japan using other injury criteria and pulses around 2010 (JNCAP, 2014).

Given the complexity of the test and with greater knowledge, Euro NCAP decided to evaluate the current protocol to see if it was still appropriate and whether it could be improved and simplified without reducing its effectiveness. The objectives of the study was to evaluate the effectiveness of the current Euro NCAP whiplash assessment procedure and current assessment criteria to determining the real-world injury outcome and whether reduced tests and output measures would equally predict real-world performance.

METHOD

Three analyses were undertaken comprising an analysis of Euro NCAP test outcome data, a logistic regression analysis, a ROC analysis, and a correlation analysis of whiplash injuries reported to Folksam Insurance, comparing crash test results and real-world injury outcome in crashes with car models tested by Euro NCAP.

Receiver Operating Characteristic curves (ROC) are a plot showing the comparative performance of systems across a varying threshold. Plotting the sensitivity and specificity of varying outcomes, in terms of true and false positives, the cumulative distribution function can then be found for each of the comparative relationships (the true positive rate against the false positive rate for the different possible cut-points of a diagnostic test). The area under the curve can then be used as a measure of the performance of the systems under test. Should one of the systems under examination score a larger area under the curve, it can be assumed to out-perform the others if the difference is statistically significant. ROC analysis is commonly used in association with a logistic regression analysis.

The Folksam material used in the correlation analysis consisted of 13,389 drivers (at least 18 years old) reporting symptoms of WAD to Folksam Insurance in Sweden between 1998 and 2013, of which 1,266 (55% females) occurred in cars tested by Euro NCAP (see Table 1). For all occupants reporting initial symptoms, the risk of permanent medical impairment was followed up according to a procedure used by Swedish insurance companies. In case an injured occupant is not recovered after approximately one year, a medical assessment is made by medical specialists to predict the impairment degree. The injured occupant is classified with a degree of medical impairment, between 1-100% depending on the injury type according to the Swedish manual for “Grading Medical Impairment” (Försäkringsförbundet 1996). All Swedish insurance companies use this manual. The manual consists of instructions of how to set the degree of medical impairment and table works for all injury types and their consequences. WAD often results in between 3-18% of medical impairment degree. The symptoms are regarded as permanent when no additional improvement in the injured patient’s mental or physical status has taken place, normally a maximum of three years after the crash.

It has been found that medical expertise in Sweden gradually has been classifying long-term consequences from whiplash associated disorders more restrictively (Kullgren et al. 2013). Therefore adjustments were made by weighting the number of occupants with long-term symptoms according to the year of impact. A reduction factor of 11% per accident year was used. In order not to change the total number of occupants with long-term symptoms the weighting was made based on accident year 2006, which is the mean accident year in the accident sample. All occupants with long-term symptoms in crashes occurring before 2006 were weighted lower and all after 2006 were weighted higher (Equation 1). By making an adjustment for accident year for each driver, the outcomes from all groups of cars under study could be compared with each other. In total 52 drivers sustained a permanent medical impairment and two out of three were females. The numbers are presented in Table 1.

$x_{pmi,adjusted} = x_{pmi}/1{,}11\text{^}(2006 - year_{accident})$, x_{pmi} = occupants with pmi Eq (1)

Table 1: Number of drivers reporting whiplash symptoms and those with pmi in cars tested by Euro NCAP.

Gender	n	n pmi
Male	547	17
Female	677	35
Unknown	42	0
Total	**1266**	**52**

Test scores according to Euro NCAP, JNCAP and IIWPG protocols were calculated based on the test performed by Euro NCAP. To increase the data available test results were also calculated based on non-official Euro NCAP test data from a test series performed before 2009. Also combinations of the three Euro NCAP pulses were calculated. For each combination or protocol, the test score was compared with the real-world outcome. The following comparisons were studied:

- Low, mid and high severity tests in Euro NCAP
- Combinations of tests (pulses) in Euro NCAP
- Individual injury criteria in each test
- Combinations of reduced criteria and reduced number of pulses

Test score expressed as per cent of maximum score versus proportion of occupants with pmi was plotted for the various combinations. For the injury criteria the abolute measured value was plotted against proportion of occupants with pmi. In each plot a linear line fit was added. The line fits were weighted for number of crashes in each point. Each point represent an average outcome in an interval. R-square values showing how well the line fits the points in each plot are calculated for each plot and p-values and 95% CI were added to each line fit.

RESULTS

The results are presented in two forms. First, a statistical analysis was performed using logistic regression analysis and Receiver Operating Characteristics (ROC) curves. This was intended to provide a rigorous statistical analysis of the overall benefit of the three existing test protocols (Euro NCAP, IIWPG, and J-NCAP) as well as show if there were differences (improvements) between the individual findings. Second, a correlational analysis was also conducted both overall as well as on components of the test protocols to indicate opportunities for future improvement.

Regression and ROC Analysis

An analysis was undertaken of the relationship between sensitivity and specificity of variations of the three test protocols that were plotted on ROC Curves. From earlier work, it was apparent that the sex and age of the victim had a substantial impact on the results in terms of risk and long term outcome, hence a regression analysis was also undertaken to control for these characteristics. The three test protocol combinations shown below were:

1. Euro NCAP comprising the median crash test (16km/h, 5.5g, with triangular pulse) and the seven test criteria (NIC, Nkm, Rebound velocity, Fx, Fz, T1 and THRC);
2. IIWPG with quantified scores, comprising a single crash pulse of 16km/h (5,5g) with triangular pulse, and four test criteria namely Time to head restraint (≤70 msec), T1 acceleration (≤ 9.5g), Fx and Fz; and
3. J-NCAP comprising a single triangular crash test pulse of 17.6km/h, and 7-injury criteria including NIC, Upper Fx (backward shear), Upper FZ (tension), Upper My (flexion), Upper My (extension), Lower Fx (backward shear) and Lower Fz (tension), and Lower My (flexion).

These data scores were all computed from Euro NCAP test data as sufficient factors are collected to create the three test protocols listed above. The ROC analysis for the 3-protocols with accompanying statistics are shown in Figure 1 and Table 2 below.

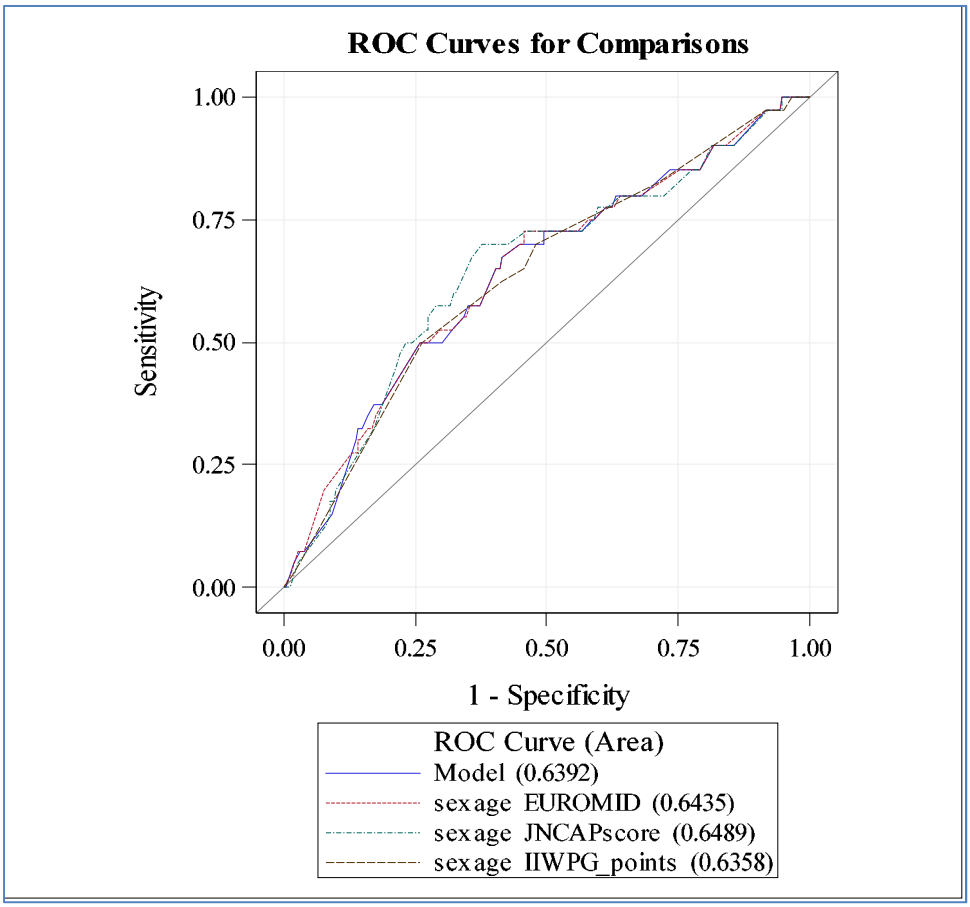


Figure 1: Comparative ROC Curves for the three test protocols

Table 2: ROC associated statistics

ROC Association Statistics							
ROC Model	**Mann-Whitney**				**Somers' D (Gini)**	**Gamma**	**Tau-a**
	Area	**Standard Error**	**95% Wald Confidence Limits**				
Model	0.6392	0.0461	0.5489	0.7295	0.2784	0.2882	0.0176
sex age EUROMID	0.6435	0.0464	0.5526	0.7344	0.2869	0.2978	0.0182
sex age JNCAPscore	0.6489	0.0462	0.5584	0.7394	0.2978	0.3084	0.0188
sex age IIWPG_points	0.6358	0.0442	0.5493	0.7224	0.2717	0.3339	0.0172

As noted above, the IIWPG subjective scale was converted into a numerical one for this comparison, which does lose a bit of information compared to the other test scores that use actual values. Nevertheless, this was necessary for the ROC analysis. The analysis showed that any apparent differences between the areas under the curves (the classic way of comparing tests) were indistinguishable from random variation. What these results do show is that any one of these tests has reasonable and statistically significant predictive value for real-world outcomes, but there was no apparent statistical difference observed between the 3-curves. It will take a considerable increase in data before apparent differences between tests will be detectable in the ROC analysis.

Correlation Analysis

Euro NCAP was significantly correlated with the other two protocols. The match between the Euro NCAP and J-NCAP appeared to be slightly stronger (less variation) than between the others as shown in Fig. 2 to 4.

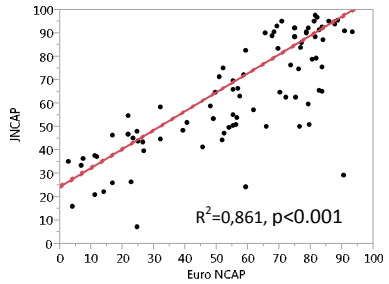


Fig 2: Euro NCAP and JNCAP scores.

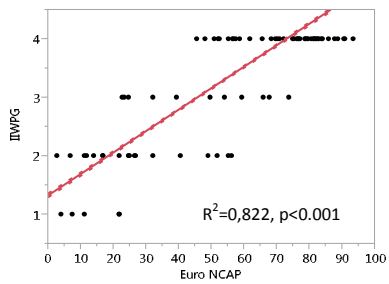


Fig 3: Euro NCAP and IIWPG scores.

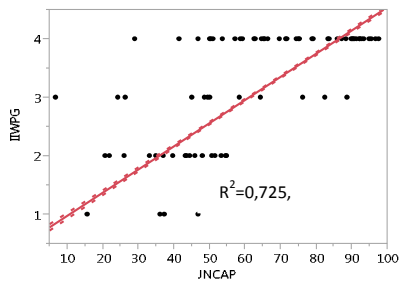


Fig 4: JNCAP and IIWPG scores.

Furthermore, all the test protocols were significantly correlated well with the risk for pmi as shown in Fig. 5 to 7.

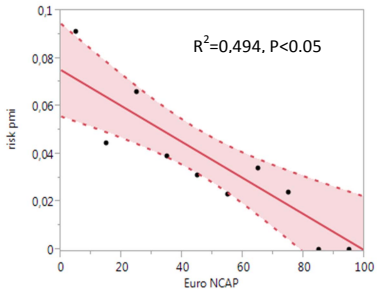


Fig 5: Risk for pmi vs Euro NCAP score.

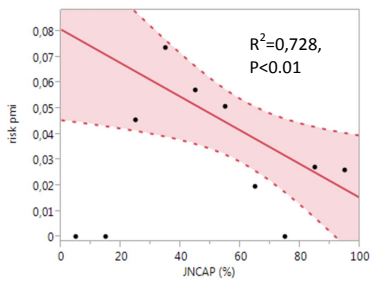


Fig 6: Risk for pmi vs JNCAP score.

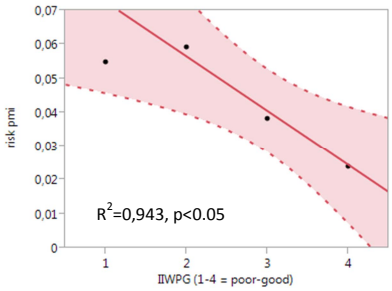


Fig 7: Risk for pmi vs IIWPG score.

Risk of Permanent Medical Impairment by Euro NCAP Crash Pulse

Fig's 8 to 10 show the association between the risk of impairment (pmi) and the test pulse for the 3-Euro NCAP tests. All three pulses in Euro NCAP correlate with risk for pmi, but the correlation for the low and mid severity pulse was stronger than for the high severity pulse.

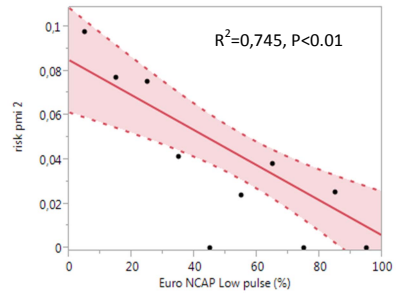


Fig 8: PMI risk vs Euro NCAP low pulse.

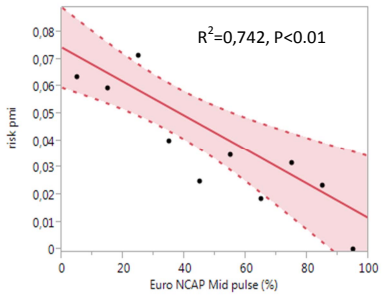


Fig 9: PMI risk vs Euro NCAP mid pulse.

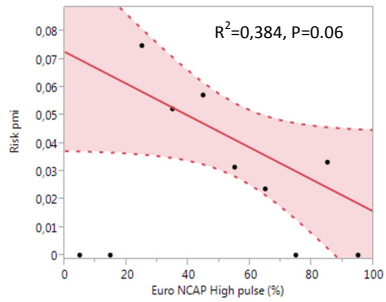


Fig 10: PMI risk vs Euro NCAP high pulse.

Further examination of the pmi risk between combinations of the three test pulses are shown in Fig's 11 to 13. The correlation between pmi risk for low and mid pulses and that for low and high pulses were significant, although the correlation between the mid and high pulse was not at the 5% probability level (around 6%).

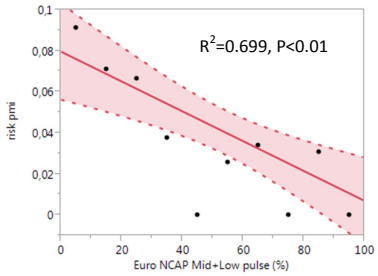


Fig 11: PMI risk vs Euro NCAP Low and Mid test pulse.

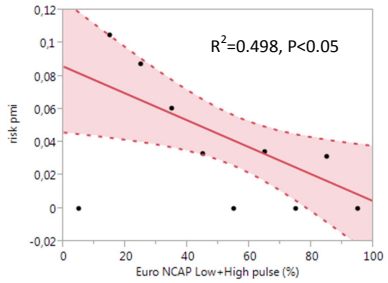


Fig 12: PMI risk vs Euro NCAP Low and High test pulse.

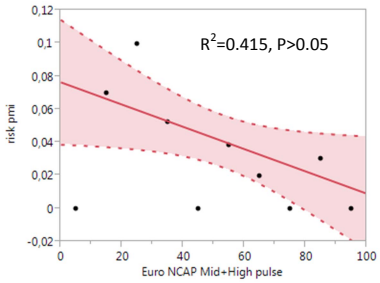


Fig 13: PMI risk vs Euro NCAP Mid and High test pulse.

Injury Criteria

Examples of results of the correlation analyses undertaken between the risk of pmi and the test criteria are shown in Fig's 14 to 19 (plots for all criteria in all three pulses can be seen in the Appedix). The criteria that show correlation ($p<0.05$) to risk of pmi were NIC in all three pulses (p just above 0.05 in the low pulse), Fz in the high pulse and rebound velocity in the low pulse. Also the full geometry assessment explained risk for pmi while geometry did not. Fx had some outliers in the low and mid pulses, an example can be seen in Fig 17, where high Fx values were measured although the risk for pmi was low. If excluding these outliers a correlation could be verified for these pulses but not for the high pulse.

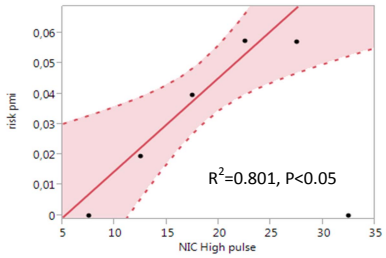


Fig14: PMI risk vs Euro NCAP NIC measure with High test pulse

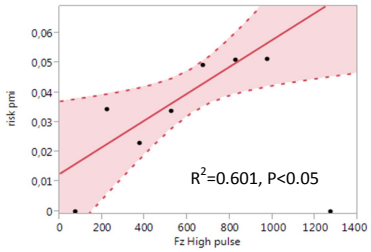


Fig15: PMI risk vs Euro NCAP Fz measure with High test pulse

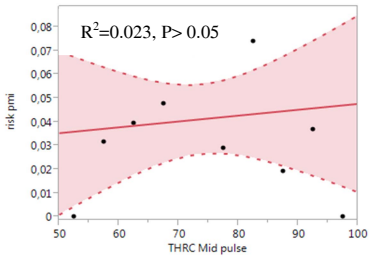


Fig16: PMI risk vs Euro NCAP THRC measure with Mid pulse

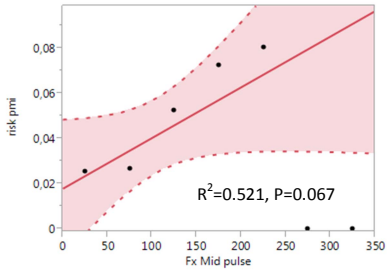


Fig17: PMI risk vs Euro NCAP Fx measure with Mid test pulse

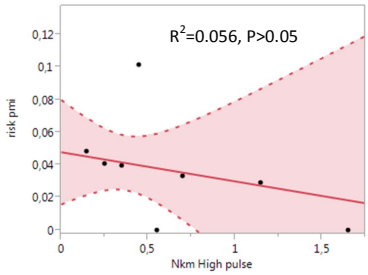


Fig18: PMI risk vs Euro NCAP Nkm measure with High test pulse

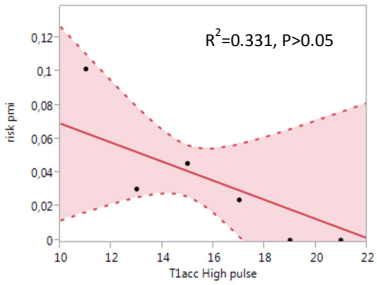


Fig19: PMI risk vs Euro NCAP T1accel measure with High test pulse

Correlations with Modified Test Pulses and Criteria

The final set of analyses examined possibilities for fewer test pulses. And based on the correlation analysis of injury criteria an attempt to using a reduced number of criteria was made using; NIC, Fx, Fz and rebound velocity. Fig's 20 to 22 show the correlations for reduced combinations of Euro NCAP test pulses and test criteria. These should be read in conjunction with Fig 4 shown earlier for the 3 test pulses and the full set of criteria. All these combinations were significant (Euro NCAP 2 pulses and all criteria had a p value just avove 0.05). However, it could not be verified which of the combinations that best explained risk for pmi.

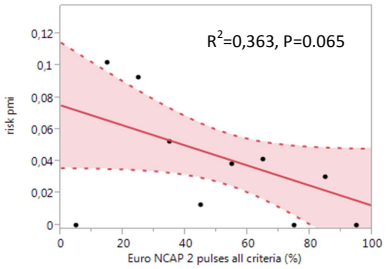


Fig 20: Two tests (Mid + High) with all criteria

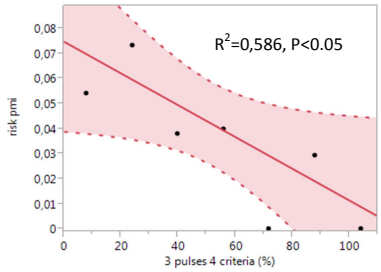


Fig 21: All tests with 4-test criteria

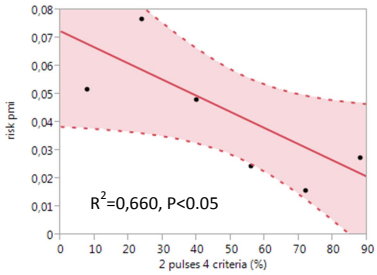


Fig 22: Two tests (Mid+High) with 4-test criteria

DISCUSSION

Efforts to reduce the number of Whiplash Associated Disorders (WAD) are essential due to the societal burden in terms of costs and suffering from impairments (Krafft 1998, Malm et al. 2008, Kullgren et al. 2013). Consumer test programmes are important tools to reduce WAD in the way they promote cars with the most effective whiplash systems. Since the first consumer test programmes were introduced by IIWPG and Folksam/SRA in 2003 and 2004 the European New Car Assessment Program (Euro NCAP) in 2009 introduced a test protocol for assessing the whiplash potential of vehicle seats as part of its vehicle safety rating system in Europe. Also Japan NCAP has introduced whiplash tests as part of the Japanese consumer test programme (JNCAP 2014).

In conducting this study, two methods were adopted in an attempt to investigate associations between the variables and the outcome criteria. The first was a rigorous parametric statistical evaluation using logistic regression and Receiver Operating Curve (ROC) analysis, while the second was a correlation analysis between aspects of the tests and whiplash outcome. While the results suggested that all protocols examined (Euro NCAP, J-NCAP and IIWPG) appeared to have a strong meaningful association with the advent of Permanent Medical Impairment (pmi), there were no statistical differences between any of the three protocols for the data available. Moreover, as noted earlier, it will require a sizable increase in the available database before this could be expected.

In conducting the correlation analysis, there were two options available for measuring real-world outcome due to the injury data available. These included (i) the likelihood of sustaining a whiplash injury leading to symptoms lasting longer than 1 month after the crash event, and/or (ii) the likelihood of sustaining on-going impairment (most often classified as a chronic outcome). Previous research by for example Davidson and Kullgren (2011) showed that while early predictions of whiplash provided greater numbers of outcomes that could be assessed, the measurement of Permanent Medical Impairment (pmi) was a more stable measure.

It was positive to find that the 3 main consumer test programmes; Euro NCAP, JNCAP and IIWPG, were equally able to predict real-world whiplash injury outcome in terms of permanent medical impairment. Previous studies have shown that both IIWPG and Folksam/SRA test programmes correlate with real-world whiplash injury outcomes (Kullgren et al. 2007, Kullgren and Krafft 2008). But this is the first time it has been shown for the Euro NCAP consumer tests. The reader needs to be careful not to infer too much from these findings though. Given the sparseness of data available so far, this can only be considered as a preliminary analysis that needs further confirmation with a much larger data set.

The current Euro NCAP whiplash test protocol includes three crash pulses of various severity and all injury criteria included in the consumer tests developed by Folksam/SRA and IIWPG in 2002 to 2003. It is a mix of 7 injury criteria and a geometrical assessment of the head rest. One of the main reasons for this is to get a robustness of the rating procedure, in the sense that the tests cover a wide range in crash severity and that many possible injury causes are covered as long as we don't know the exact cause of the whiplash symptoms. Previous studies have shown correlations between various of the injury criteria used and real-world injury outcome (see for example Davidsson and Kullgren 2011). The results reported here showed that some of the criteria used were less able to predict with real-wold whiplash injury outcome. In particular, time to head rest contact (THRC) and the acceleration of the T1 vertebrae (T1acc) were not correlated with Permanent Medical Impairment as were Nkm and Fx (although Fx was significant however for the Mid pulse, although not for the Low and High pulses). In addition, for the car models with the highest measured Fx values the risk for pmi was very low. Note that the numbers of observations for these points in the plots involved only one or two crashes.

Even though the accident material covered 16 years (1998 to 2013) of crashes reported to Folksam Insurance, the number of available cases was too low to provide statistically robust results. Nevertheless, the findings reported here for reductions in the number of tests and outcome criteria were promising, suggesting possibile options for explaining the risk for pmi and as a basis for potentially developing the Euro NCAP protocol in the years ahead.

Limitations

By far the biggest limitation experienced with this analysis was in the number of suitable cases of Permanent Medical Impairment in the current Folksam database. While this could be overcome to some degree by using the outcome of whiplash injury with symptoms lasting longer than 1 month after the crash event, from previous studies, this may be offset by the degree of variation in the findings.

To increase the data avalable for analysis it was necessary to use non published Euro NCAP data nased on tests performed before 2009. It was not fully verified if these tests fully met all Euro NCAP protocol requeirements.

The ROC analysis is well-established as a means to compare the performance of test criteria against a gold-standard outcome measure (here, real-world injury). Substantially more data would be required before the tests studied can be distinguished using these criteria.

SUMMARY AND CONCLUSIONS

The analyses conducted here were aimed at identifying real-world whiplash associations with existing test protocols used by Euro NCAP, J-NCAP, and IIWPG. Given the paucity in the data available, the finding here should be regarded as preliminary findings at this stage. Of interest, Euro NCAP, JNCAP and IIWPG were all found to be significanly correlated with each other and correlated to some degree with the risk for WAD leading to permanent medical impairment. There were signs that that there could be refinements in the number of test criteria. There was a suggestion that reductions in both the number of tests and criteria could still provide significant associations with Permanent Medical Impairment, but that further research is warranted to further test its robustness.

REFERENCES

Davidsson J, Kullgren A (2011). Evaluation of Seat Performance Criteria for Rear-end Impact Testing. 22th International Conference on the Enhanced Safety Vehicles (ESV), Washington, DC, National Highway Traffic Safety Administration.

Eichberger A, Geigl B C, Moser A, Fachbach B, Steffan H, Hell W, Langwieder C (1996) Comparison of different car seats regarding head neck kinematics of volunteers during rear-end impact. Proc. of the Int. IRCOBI Conf. on the Biomechanics of Injury. Dublin, Ireland. pp153-164.

Farmer C, Zuby D, Well J, Hellinga L (2008). "Relationship of dynamic seat ratings to real-world neck injury rates." Traffic Inj Prev 9(6): 561-567.

Försäkringsförbundet (1996) Gadering av medicinsk invaliditet –96 (only in Swedish). IFU Utbildning AB, Stockholm Sweden.

Jakobsson L (1998). Automobile Design and Whiplash Prevention. Whiplash Injuries: Current Concepts in Prevention, Diagnoses and Treatment of the Cervical Whiplash Syndrome. R. Gunzberg and M. Szpalski. Philadelphia, Lippincott-Raven Publishers: 299-306.

Jakobsson L, Norin H, Svensson MY (2004) Parameters Influencing AIS 1 Neck Injury Outcome in Frontal Impacts, Traffic Injury Prevention, Vol. 5, No. 2, pp. 156–163.

Jakobsson L. Whiplash Associated Disorders in Frontal and Rear-End Car Impacts. Biomechanical. Thesis for the degree o doctor of philosophy. Crash Safety Dividson , Dep of Machine and Vehicle Systems. Chalmers University of Technology, Sweden 2004. Jakobsson et al. 2008

JNCAP (2014). Car Safety Performance Guidelines; New Car Assessments 2014.3, http://www.nasva.go.jp/mamoru/en/download/car_download.html

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. Karolinska Institutet, Institution of Clinical Neuroscience, Section for Personal Injury Prevention Stockholm, Sweden,

Krafft M, Kullgren A, Lie A, Tingvall C (2003) The Risk of Whiplash Injury in the Rear Seat Compared to the Front Seat in Rear Impacts, Traffic Injury Prevention, Vol. 4, No. 2, pp. 136–140

Krafft M, Kullgren A, Lie A, Tingvall C (2004). Assessment of Whiplash Protection in Rear Impacts - Crash Tests and Real-life Crashes. Test report of the Folksam/SRA consumer test program 2004. Folksam 10660 Stockholm, Sweden.

Eriksson L, Boström O, Krafft M (2003) Validation of neck injury criteria using reconstructed real-life rear-end crashes with recorded crash pulses. Proc. Of the 18 th Int. Techn. Conf. on ESV, paper 344.

Kullgren A, Krafft M, Lie A, Tingvall C (2007). The Effect of Whiplash Protection Systems in Real-Life Crashes and their Correlation to Consumer Crash Test Programmes. 20th International Technical Conference on the Enhanced Safety of Vehicles (ESV), Lyon, France.

Kullgren A, Krafft M (2008) Whiplash research in Sweden: results from real-life crashes. EVU conference, Nice

Kullgren A, Stigson H, Krafft M, (2013). Development of Whiplash Associated Disorders for Male and Female Car Occupants in Cars Launched Since the 80s in Different Impact Directions. Int. IRCOBI Conf. on the Biomechanics of Injury, Gothenburg, Sweden.

Lundell B, Jakobson L, Alfredsson B, Lindström M, Simonsson L (1998) The WHIPS seat – A car seat for improved protection against neck injuries in rear end impacts. Paper No 98-S7-O-08, Proc. 16th ESV Conf, 1998, pp. 1586-1596.

Malm S, Krafft M, Kullgren A, Ydenius A, Tingvall C. (2008). "Risk of permanent medical impairment (RPMI) in road traffic accidents." Annu Proc Assoc Adv Automot Med 52: 93-100.Proceedings of IRCOBI Conference, Berlin, Germany.

Sekizuka M (1998) Seat Designs for Whiplash Injury Lessening, Proc. 16th Int. Techn. Conf. on ESV, Windsor, Canada.

Thatcham (2013) www.thatcham.org

Viano D and Olsen S (2001) The Effectiveness of Active Head Restraint in Preventing Whiplash. Journal of Trauma 2001; Vol.51: 959-969

Watanabe Y, Ichikawa H, Kayama O, Ono, K, Kaneoka K, Inami S (2000) Influence of Seat Characteristics on Occupant Motion in Low-Velocity Rear-End Impacts, Accid. Anal. Prev. 32(2): 243–250.

Wiklund K, Larsson H (1997) SAAB Active Head Restraint (SAHR) - Seat Design to Reduce the Risk of Neck Injuries in Rear Impacts, SAE Paper 980297, Warrendale.

APPENDIX

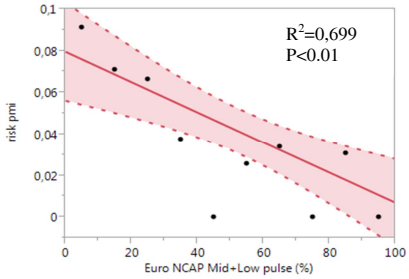


Fig A1. Risk for pmi vs low+mid pulses.

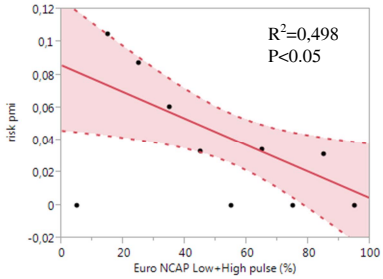


Fig A2. Risk for pmi vs low+high pulses.

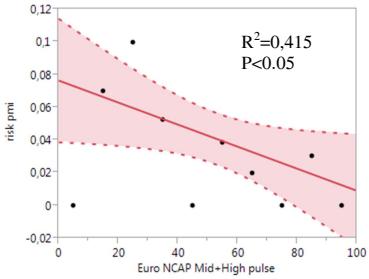


Fig A3. Risk for pmi vs mid+high pulses.

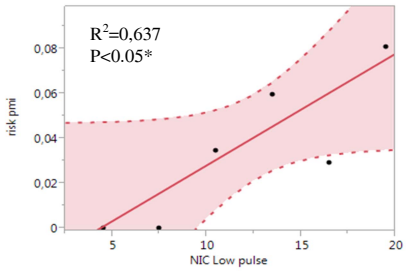


Fig A4. Risk for pmi vs NIC low pulse.

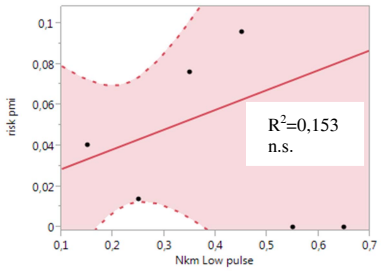


Fig A5. Risk for pmi vs Nkm low pulse.

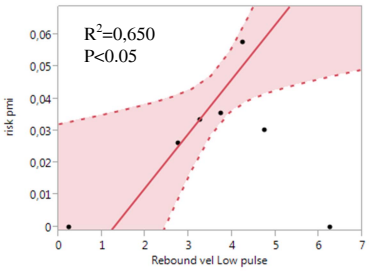


Fig A6. Risk for pmi vs reb vel low pulse.

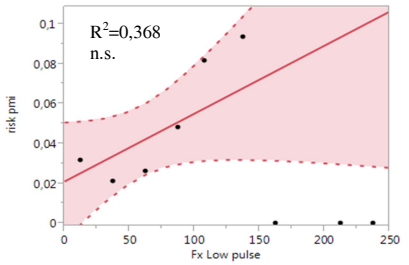


Fig A7. Risk for pmi vs Fx low pulse.

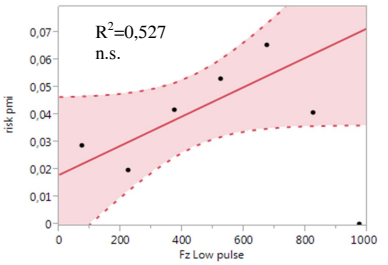


Figure A8. Risk for pmi vs Fz low pulse.

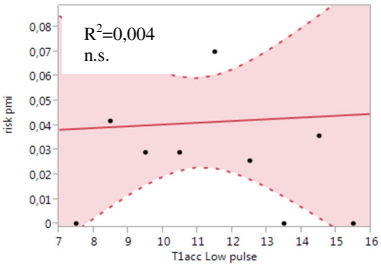


Fig A9. Risk for pmi vs T1acc low pulse.

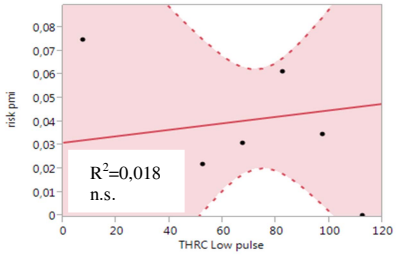


Fig A10. Risk for pmi vs THRC low pulse.

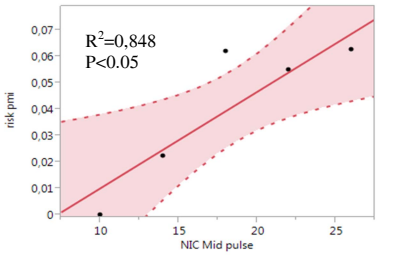


Fig A11. Risk for pmi vs NIC mid pulse.

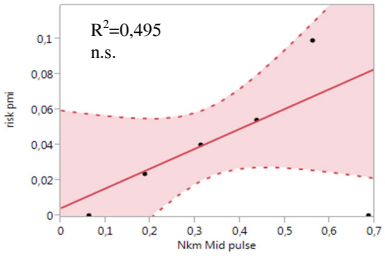


Figure A12. Risk for pmi vs Nkm mid pulse.

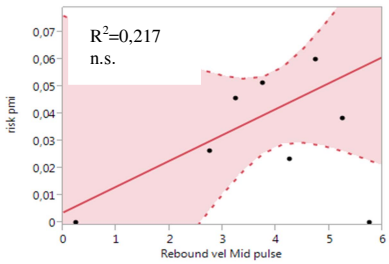


Fig A13. Risk for pmi vs reb vel mid pulse.

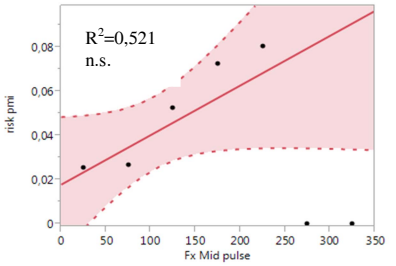


Fig A13. Risk for pmi vs Fx (mid pulse).

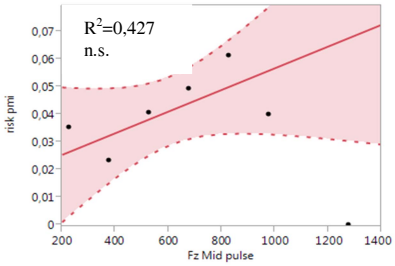


Fig A14. Risk for pmi vs Fz mid pulse.

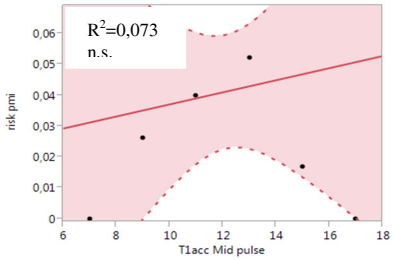


Fig A15. Risk for pmi vs T1 acc mid pulse.

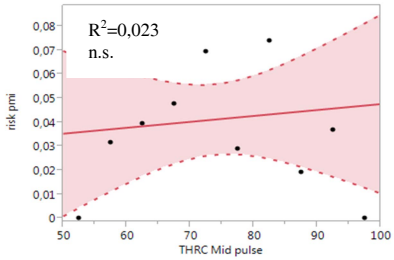


Fig A16. Risk for pmi vs THRC mid pulse.

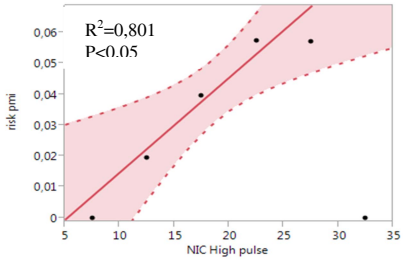


Fig A17. Risk for pmi vs NIC high pulse.

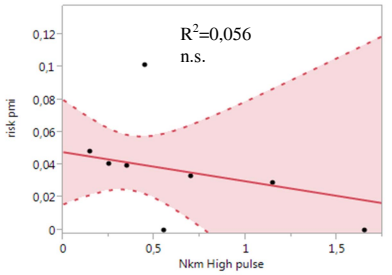


Fig A18. Risk for pmi vs Nkm high pulse.

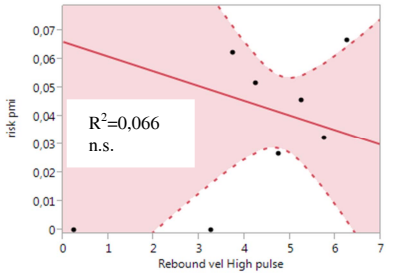


Fig A19. Risk for pmi vs reb vel high pulse.

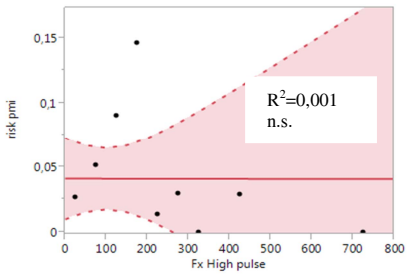


Fig A20. Risk for pmi vs Fx high pulse.

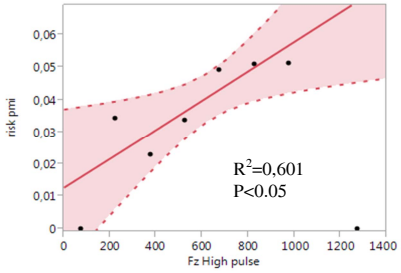


Fig A21. Risk for pmi vs Fz high pulse.

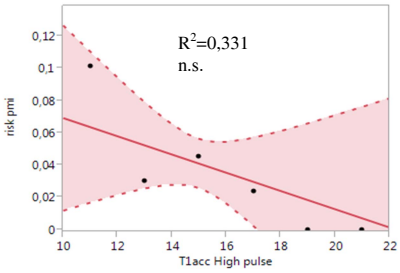


Fig A22. Risk for pmi vs T1acc high pulse.

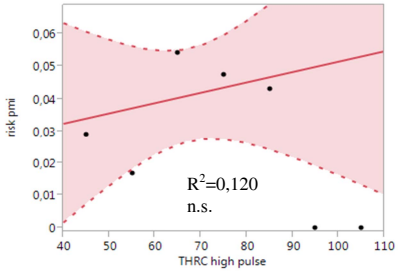


Fig A23. Risk for pmi vs THRC (high pulse).

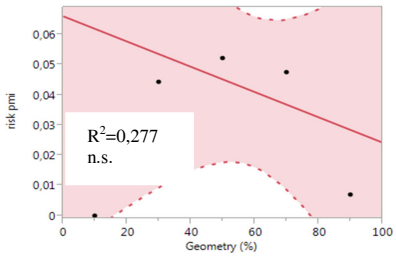


Fig A24. Ris for pmi vs geometry.

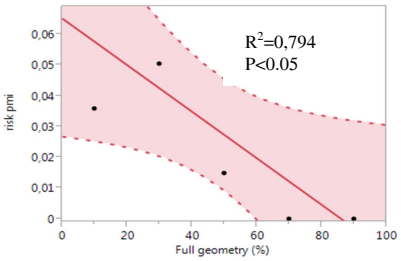


Fig A25. Risk for pmi vs full geometry.

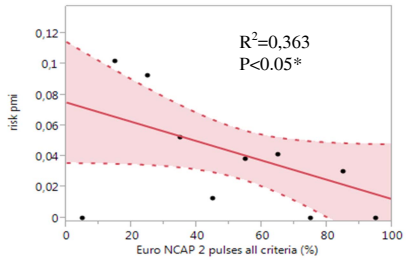


Fig A26. Risk for pmi vs Euro NCAP 2 pulses (mid and high) and all criteria.

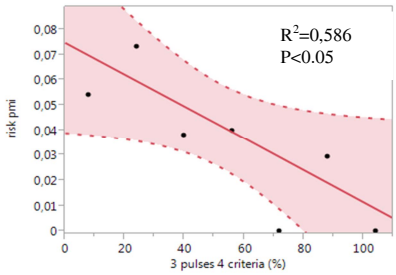


Fig A27. Risk for pmi vs all Euro NCAP pulse and 4 criteria; NIC, Fx, Fz and reb vel.

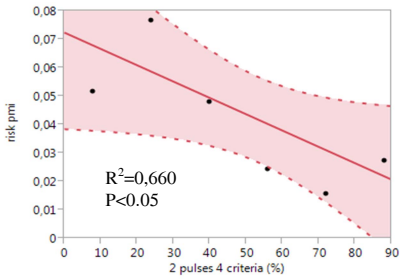


Fig A28. Risk for pmi vs Euro NCAP 2 pulses mid and high and 4 criteria; NIC, Fx, Fz and reb vel.

A METHODOLOGY TO DERIVE PRECISION REQUIREMENTS FOR AUTOMATIC EMERGENCY BRAKING (AEB) TEST PROCEDURES

Patrick Seiniger, Jost Gail

Bundesanstalt für Straßenwesen (Federal Highway Research Institute) Brüderstraße 53, 51427 Bergisch Gladbach, Germany

Phone: +49 2204 43619
Fax: +49 2204 43676
E-mail: seiniger@bast.de

Paper ID#: 15-0315

ABSTRACT

AEB Systems are becoming important to increase traffic safety. Test procedures in testing for consumer information, manufacturer self-certification and technical regulations are used to ensure a certain minimum performance of these systems. Consequently, test robustness, test efficiency and finally test cost become increasingly important.
The key driver for testing effort and test costs is the required repeatable accuracy in a test design - the higher the accuracy, the higher effort and test costs. On the other hand, the performance of active safety systems depends on time discretization in the environment perception and other sub-systems: for instance, typical sensors supply information with a cycle time of 50 - 150 ms. Time discretization results in an inherent spread of system performance, even if the test conditions are perfectly equal.
The proposed paper shows a methodology to derive requirements for a test setup (e.g. test repeats, use of driving robots, ...) as function of AEB system generation and rating method (e.g. Euro NCAP points awarded, pass/fail, ...). While the methodology itself is applicable to AEB pedestrian and AEB Car-Car scenarios, due to the lack of sufficient test data for AEB Car-Car, the focus of this paper is on AEB pedestrian scenarios.
A simulation model for the performance of AEB Pedestrian systems allows for the systematic variation of the discretization time as well as test condition accuracy. This model is calibrated with test results of 4 production vehicles for AEB Pedestrian, all fully tested by BASt according to current Euro NCAP test protocols.
Selected parameters to observe the accuracy of the test setup in case of pedestrian AEB is the calculated impact position of pedestrian on the vehicle front (as if no braking would have occurred), and the test vehicle speed accuracy. These variable was shown in real tests to be repeatable in the range of ± 5 cm and ± 0,25 km/h, respectively, with a fully robotized state of the art test setup.
The sensitivity of AEB performance (measured in achieved speed reduction as well as overall rating result according to current Euro NCAP rating methods) towards discretization and the sensitivity of performance towards test accuracy then is compared to identify economic yet robust test concepts.
These comparisons show that the available repeatability accuracy of current test setups is more than sufficient for today's AEB system capabilities. Time discretization problems dominate the performance spread especially in test scenarios with a limited pedestrian dummy reveal time (e.g. child behind obstruction, running adult scenarios with low car speeds). This would allow to increase test tolerances to decrease test cost.
A methodology which allows to derive the required tolerances in active safety tests might be valuable especially for NCAPs of emerging countries that do not have the necessary equipment (e.g. driving robots, positioning units) available for the full-scale and high tolerance EuroNCAP active safety procedures yet still want to rate active safety systems, thus improving the global safety.

INTRODUCTION

AEB Systems are becoming important to increase traffic safety. Test procedures in testing for consumer information, manufacturer self-certification and technical regulations are used to ensure a certain minimum performance of these systems. Consequently, test robustness, test efficiency and finally test cost become increasingly important.
One of the major discussion points during the implementation of AEB test procedures was the required precision of test execution. The initial scenarios (as implemented in 2014) had been longitudinal scenarios. Both vehicles (target as well as vehicle under test) move on parallel - ideally identical – tracks. The main parameters which drive test cost are the precision of both vehicles' speeds, and the lateral displacement between the vehicles at all times.

Things become more complicated for scenarios where the target brakes during the test conduction: the initial longitudinal displacement between both vehicles should ideally be constant, and the target deceleration process, characterized by brake ramp and target deceleration, needs to be repeatable.
Finally, lateral traffic accident situations such as the proposed pedestrian scenarios require an exact match between positions of both target and vehicle under test during the constant speed phase of the test execution; otherwise the AEB system is presented with a different situation. A parameter representing this match of positions is the theoretical impact position of the target on the vehicle-under-test's front, calculated for a situation without automatic braking.
Testing always introduces uncertainties into how good these parameters can be maintained. The more complex the equipment is, the lower become the tolerances, and the higher is the precision of testing. Goal of this paper is to identify the relation between tolerances and expected certainty of results (e.g. speed reduction for the pedestrian AEB systems).

Research Questions and Methodology

The goal of this paper is to identify the relation between precision in test execution and variability in test results for AEB pedestrian test setups. Variability is influenced by test execution precision, but also by limitations inside current AEB systems.

- How does the test execution precision in AEB pedestrian tests influence the variability in speed reduction and overall rating?
- How do technical limitations in pedestrian AEB systems influence the variability in speed reductions and overall rating?

These questions will be answered with a simulation model for AEB pedestrian tests. The model allows to introduce tolerances in the test execution and as well as technical limitations (e.g. sensor angles, processing times, time discretization). A model for AEB Car-Car braking scenarios is developed as well, however there is not sufficient data for model validation.
Model input besides the mentioned tolerances and technical simulations is the test condition, output is the achieved speed reduction This model is parameterized and validated using all possible test data. Validation is successful if the simulation results match the test track results sufficiently well. Sufficiently well in this case means that the observed effects (e.g. variations, peak avoidance speeds, tendencies of speed decrease) can be replicated.
Results of the sensitivity investigations then are presented as plots overall rating variation as function of the varied parameter.

SIMULATION MODEL

The simulation model is loosely based on the model presented in [Seiniger, 2013]. It simulates the braking process of a vehicle, including brake delay time, deceleration ramp and deceleration levels. In [Seiniger, 2013], the model determined the brake time using certain paradigms (e.g. when the accident cannot be avoided by driver or pedestrian). The updated model now determines the appropriate brake time using AEB logic, sensor and visibility characteristics. These parameters had been derived from test data, gained within the AsPeCSS project [ASPECSS, 2014] and in BASt tests for manufacturers. Brake time determination differs whether the simulated test is an AEB pedestrian test or an AEB Car-Car test.

Generic Simulation Model Structure

Main part of the simulation model is the calculation of speed reduction as function of brake parameters, brake time and vehicle initial speed. The brake time is determined either by brake strategy (in cases where target detection is not a limiting factor) or by the remaining time before the impact after sensor detection (in other cases). The target reveal time (only applicable for pedestrian AEB simulation) is not only limited due to obstructions, but in some situations also by the sensor field-of-view and / or additional processing time. The overall structure of the simulation model is shown in Figure 1.

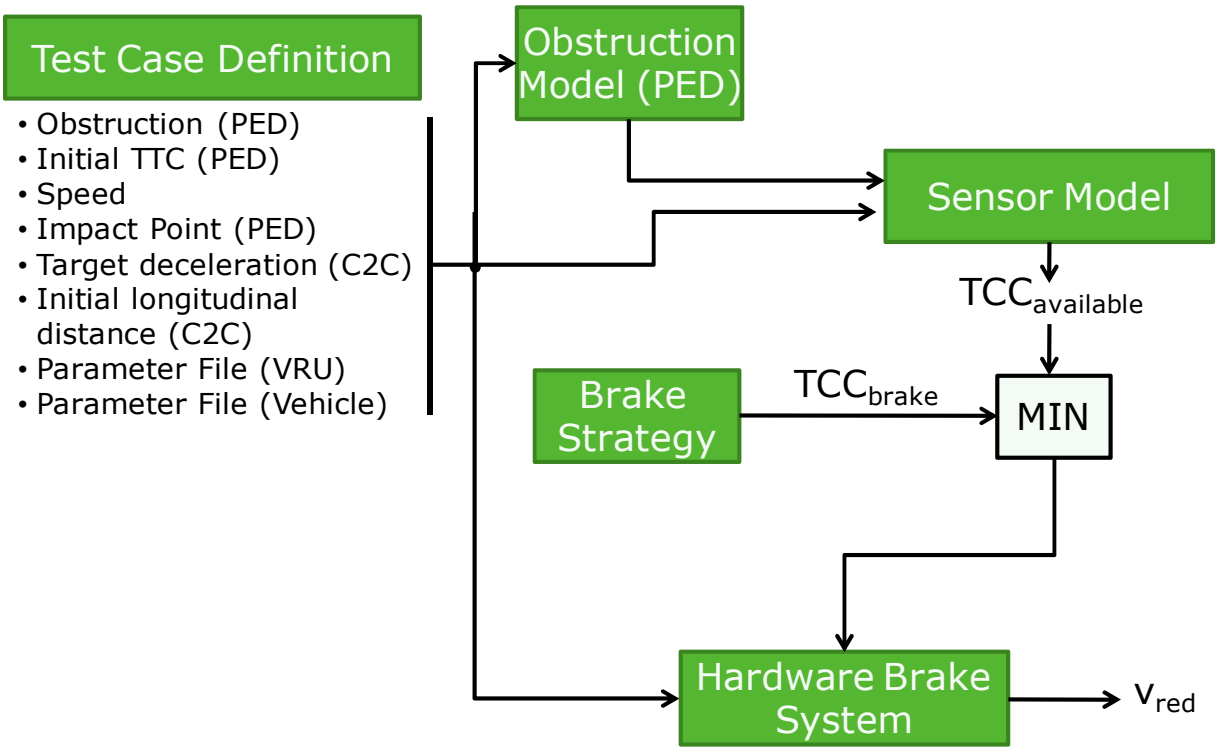


Figure 1. **General structure of the simulation model (PED = pedestrian, C2C = Car-Car, MIN = minimum value of $TTC_{available}$ and TTC_{brake} is selected)**

Sensor and obstruction models for Pedestrian AEB tests

The concept behind the obstruction and sensors model block in case of pedestrian AEB is depictured in Figure 2. For Field-of-view investigation, the angular position of the target towards the sensor is evaluated. The sensor model calculates the time at which the target enters the sensor field of view, if this is a limiting factor, and the obstruction model calculates the time at which the target appears from behind the obstruction, if this is a limiting factor. In both cases, the percentage of the target that is required for proper detection is a parameter of the simulation. Other parameters within the obstruction and sensor model block are the size and position of the obstruction with respect to target and vehicle-under-test, the sensor field-of-view angle, the sensor longitudinal position with respect to the vehicle front, and the required sensor processing time, after the target becomes visible.

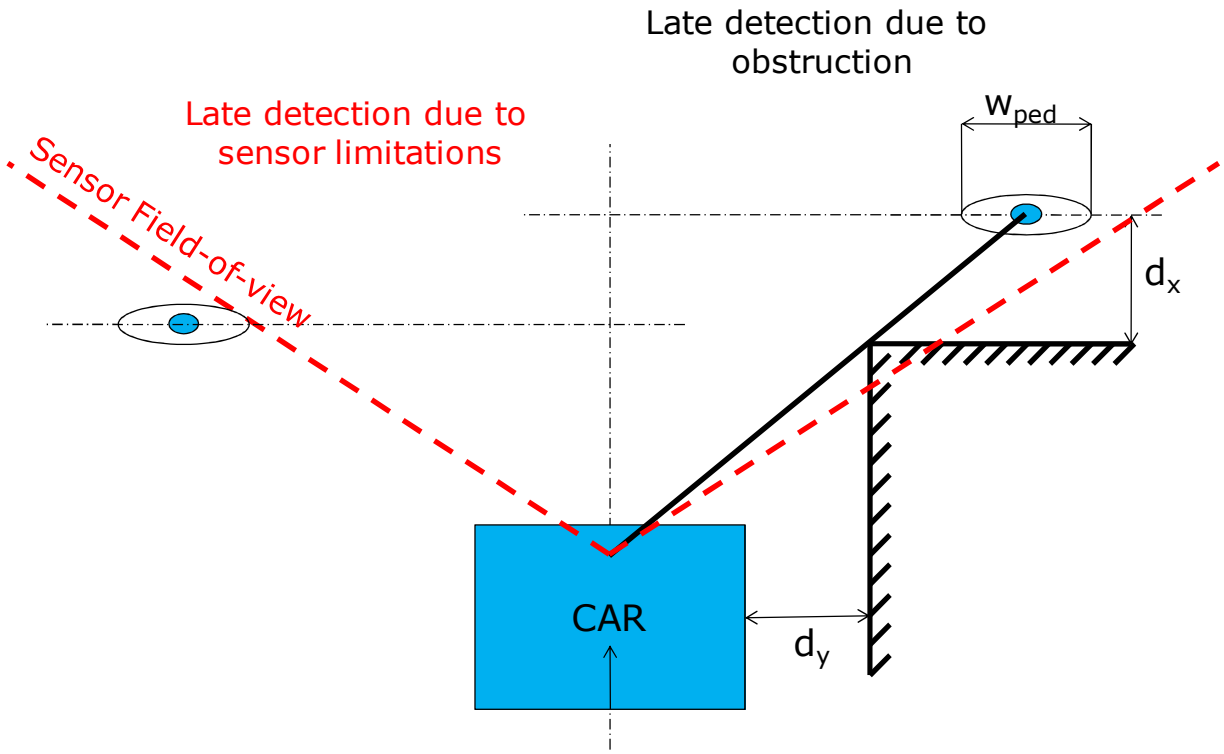


Figure 2. **Concepts behind the sensor model and obstruction model blocks. Variables are the lateral position of obstruction with respect to the vehicle d_y, the longitudinal position of the obstruction with respect to the pedestrian target d_x, the size of the pedestrian target w_{ped}.**

An important parameter that is responsible for variations in measured speed reduction even with perfectly constant test parameters is the sensor sample rate. All sensors sample their data at discrete times, and usually the interval between sampling times is constant. Assuming that the AEB system brake decision requires at least one measurement (=sample) to verify that brake conditions exist, this final verification is available with a maximum variability equal to the sample rate, see Figure 3.

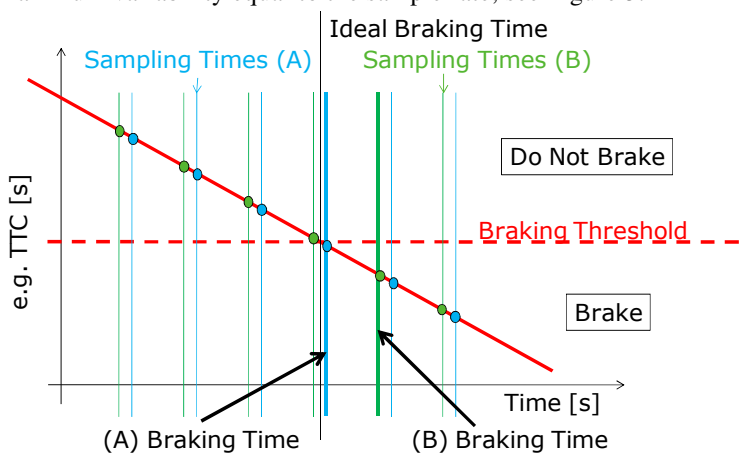


Figure 3. **Sensor time discretization creates brake timing variations: an example using TTC as threshold for brake timing. The ideal brake timing would be at the point where TTC crosses the threshold. In case A, blue, a new measurement (dots on the TTC line, accompanied by vertical lines) is available just after TTC crosses the threshold, so almost no additional time delay due to sensor time discretization is introduced. In case B, green, a measurement occured just before the TTC crossed the threshold, so the new measurement is**

available almost one complete sample time after TTC has crossed the threshold - and thus, almost one complete sample time is introduced as the additional time delay. Since the sampling process starts at system start far before the test, the probability for any time delay between no delay and full sampling time should be equal.

Brake Strategy for Pedestrian AEB Systems

Identification of brake strategy requires some system identification from the available test data. Various strategies are used by the in total 5 cars that were available for validation. An appropriate criterion to characterize the threat of collision is the value Time-To-Collision TTC [Winner, 2011]. TTC is the quotient of relative distance in direction of travel Δx and relative speed Δv:

$$TTC = \frac{\Delta x}{\Delta v}.$$

The variable TTC describes the current time that is left until a collision occurs if speeds of target and vehicle-under-test do not change.
Test data shows that all vehicles use a threshold for the lateral position of the target, a threshold for the Time-To-Collision value (TTC), or both, and in some cases, different thresholds for different pedestrian speeds have been observed. Note that a collision will occur in all of the defined test cases, if no braking is initiated.

Brake Strategy for Car-Car AEB Systems

The most relevant and demanding Car-Car test scenario with respect to required accuracy is the braking scenario: two vehicles initially travel at an equal and constant speed, with a constant longitudinal distance, until the target vehicle decelerates. The target vehicles' deceleration level and the time until this fully developed deceleration is achieved are test parameters. In this case, the variable TTC as described above cannot be used, since it does not take the target deceleration into account. A more appropriate variable is the "enhanced TTC" ETTC [Winner, 2011]:

$$ETTC = \frac{\sqrt{\Delta v^2 + 2 \cdot D_{rel} \cdot \Delta x} - \Delta v}{D_{rel}},$$

introducing the relative deceleration between both vehicles D_{rel}. Test data shows that ETTC characterizes the brake timing relatively well: start of braking consistently occurs at a constant ETTC threshold (only test data from one vehicle test series used).
The sampling time effect on time variability should be relevant for these tests as well, but sensor detection usually is not a limiting factor since the target vehicle is in full sensor field of view during the whole experiment.

Hardware Brake System

The model for the hardware brake system calculates the achieved speed reduction for a given TTC of commanded braking, relative velocity and initial conditions, taking a constant delay time after brake command, constant brake jerk and maximum deceleration into account. All calculations are numerical.
Additionally to calculation of speed reduction, this block also evaluates whether the accident is avoided or not: for pedestrian AEB tests, in some cases pedestrians gain enough time to clear the path due to the brake intervention, and these situations then are counted as avoided accidents.

Model Parameter List

Table 1 shows a full set of the available model parameters for AEB pedestrian as well as AEB Car-Car.

Table 1. **Simulation model parameter list**

Parameter	Category	Description	Typical value
TTC_AEB_DESIRED	Brake Strategy	TTC Threshold for brake onset	1 s
DIST_PEDESTRIAN	Brake Strategy	Lateral distance threshold between pedestrian and vehicle center at brake onset	2.2 m

ADJUST_DIST_RUNNING	Brake Strategy	Increase DIST_PEDESTRIAN by this factor for running pedestrians (8 km/h rather than 5 km/h)	1.2
ETTC_BRAKE	Brake Strategy	ETTC Threshold for brake onset in Car-Car scenarios	1
TTC_DELAY	Sensor Characteristics	Time that is required from first sensor measurement until proper target recognition	0.2 s
ADJUST_T_DELAY_SPEED	Sensor Characteristics	Higher relative speeds will increase the detection time, because optical flux / differences between two images become to great.	YES (for camera)
T_DELAY_SPEED	Sensor Characteristics	Speed threshold after which detection time increases	35 km/h
T_DELAY_HIGH_SPEED	Sensor Characteristics	Factor for detection time increase	2
ADJUST_T_DELAY_DISTANCE	Sensor Characteristics	Higher distance will increase the detection time for stereo cameras, because stereo effect vanishes for far objects	YES (stereo camera)
T_DELAY_DISTANCE	Sensor Characteristics	Distance	15 m
T_DELAY_HIGH_DISTANCE	Sensor Characteristics	Factor for detection time increase at high distances	2
SENSOR_ANGLE	Sensor Characteristics	Field of view-angle of the most relevant sensor	20°
TARGET_PROPORTION	Sensor Characteristics	What proportion of the target needs to be in sight for proper target recognition	50%
SENSOR_X_POS	Sensor Characteristics	Most relevant sensor longitudinal position with respect to vehicle front	0 m (for RADAR)
FRAMERATE	Sensor Characteristics	Most relevant sensor frame rate	10 Hz
T_DEAD_AEB	Brake Hardware	Delay between brake command and brake onset	0,35 s
JERK_AEB	Brake Hardware	Maximum rate of deceleration increase	20 m/s^3
A_MAX_AEB	Brake Hardware	Maximum achievable deceleration	6 m/s^2
W_VEHICLE	Vehicle	Width of vehicle	1.85 m
W_PED	Target	Width of target (e.g. length of bicycle target, width of pedestrian dummy)	0.6 m (Pedestrian Dummy)

MODEL VALIDATION

AEB for Pedestrians

The quality of the simulation model can be estimated by comparing simulation results with test data. For the validation process of the pedestrian AEB model, test data from four different vehicles is available. All four vehicles have an own parameter set. Parameters have been identified from the test data as well.
Simulations and test results for four vehicles and four scenarios are shown in Figure 4.

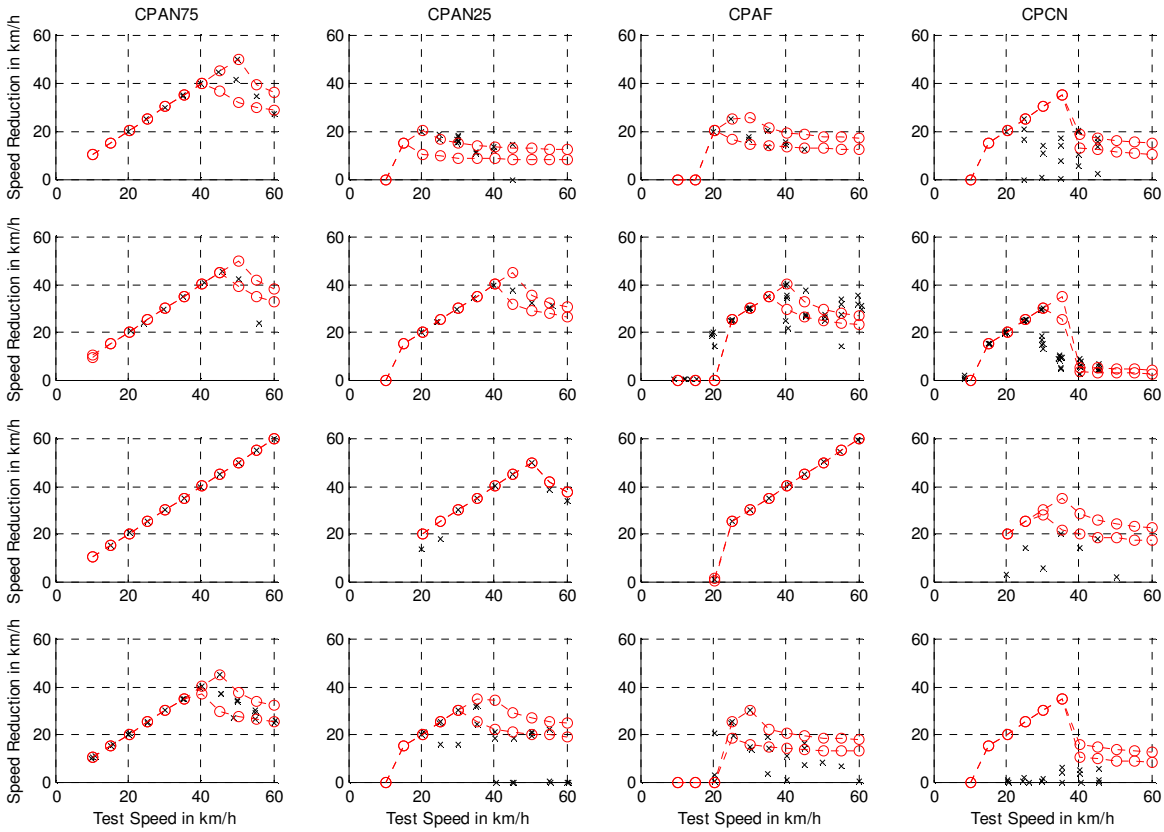


Figure 4. **Comparison of measured speed reductions (black crosses) and simulation results (red circles) for pedestrian AEB tests. Scenarios according to Euro NCAP pedestrian scenarios, CPAN75=Car-Pedestrian-Adult-Near Side-75%, CPAN25=Car-Pedestrian-Adult-Near Side-25%, CPAF=Car-Pedestrian-Adult-Far Side, CPCN=Car-Pedestrian-Child-Nearside. Two different plots for simulation results are available: one for no time delay due to sensor sampling time (in general the upper line), and one for full time delay due to sensor sampling time.**

Simulation results match test data sufficiently well: they reproduce the general tendencies of the test data like peak avoidance speeds, minimum activation speeds, and result variability. This leads to the conclusion that the model structure is valid and the model can be used for the projected sensitivity studies.

SENSITIVITY ANALYSIS

Methodology for sensitivity analysis

The goal of this paper is to identify the influence of tolerances in test execution and variability in technical systems on the test result in AEB tests. A simulation model with parameter sets for four different real world vehicles is used to answer these questions. The baseline is ideal test execution according to current Euro NCAP scenarios for pedestrian AEB and a sensor processing factor of 0.5 (meaning additional delay of half the sensor sampling time). Based on this, selected parameters (one at a time) are modified to represent inaccuracies in test setup. The results are given in one combined plot per vehicle and scenario, resulting in an array of 16 plots.

These 16 plots in Figure 5 show much detail about the required test precision, but for an overview, an integral criterion is needed: Euro NCAP's assessment method is such an appropriate integral criterion.
Euro NCAP awards different points per test speed, corresponding to the real world relevance of that specific speed (for details, see [Seiniger 2015]). Full points for lower speeds are awarded for full speed reduction using a sliding scale. Full points for higher speeds (>45 km/h) are available if the speed reduction at the individual test is more than 20 km/h, and the test is conducted only if the preceding test was passed. The scoring is shown Table 2.

Table 2. **Rating parameters.**

Index "k"	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Test speed	20	25	30	35	40	45	50	55	60
Points	1	2	2	3	3	3	2	1	1
Sliding Scale	1	1	1	1	1	0	0	0	0
PassFail	0	0	0	0	0	1	1	1	1
Method	• Sliding scale: ratio of speed reduction and initial speed					• Pass for speed reduction > 20 km/h • Test only executed if previous test speed is passed			

In a mathematical formulation, the rating result per scenario will become:

$$Rating = \sum_{k=1}^{k=n} \mathrm{SC}(k) \cdot \mathrm{Points}(k) \cdot \frac{v_{\mathrm{x,red,k}}}{v_{\mathrm{x,0,k}}} + \mathrm{PF}(k) \cdot \mathrm{Points}(k) \cdot < v_{\mathrm{x,red,k}} - v_{\mathrm{x,red,min}} >^{0} ,$$

with the vectors "Sliding scale SC", "PassFail PF", "Points" and "$v_{x,red,min}$ = minimum speed reduction = 20 km/h" as defined in Table 2, and using the Föppl-parenthesis "$< \ >^{0}$" which becomes 0 for results less than zero and 1 elsewise.

Results

Test case definition, color coding for the plots and rating results relative to each baseline condition per vehicle is shown in Table 3.

Table 3. **Parameter variations, coding as shown in *Figure 5* and overall rating results compared to baseline in %, for vehicles A to D**

Variant No.	Variation	Color	Marker	Rating A	B	C	D
1	Baseline	Black	.	100%	100%	100%	100 %
2	Earlier Detection, no influence of sampling time	Red	+	107%	101%	104%	104 %
3	Later Detection, full influence of sampling time	Red	O	93%	99%	98%	90%
4	Impact Point 5 cm towards Dummy	Blue	O	96%	99%	99%	95%
5	Impact Point 20 cm towards Dummy	Green	O	84%	91%	99%	75%
6	Impact Point 5 cm away from Dummy	Blue	+	104%	101%	101%	102 %
7	Impact Point 20 cm away from Dummy	Green	+	133%	101%	102%	120 %
8	Test speed 1 km/h below nominal speed	Cyan	+	104%	101%	105%	105 %
9	Test speed 1 km/h above nominal speed	Cyan	O	98%	96%	102%	101 %

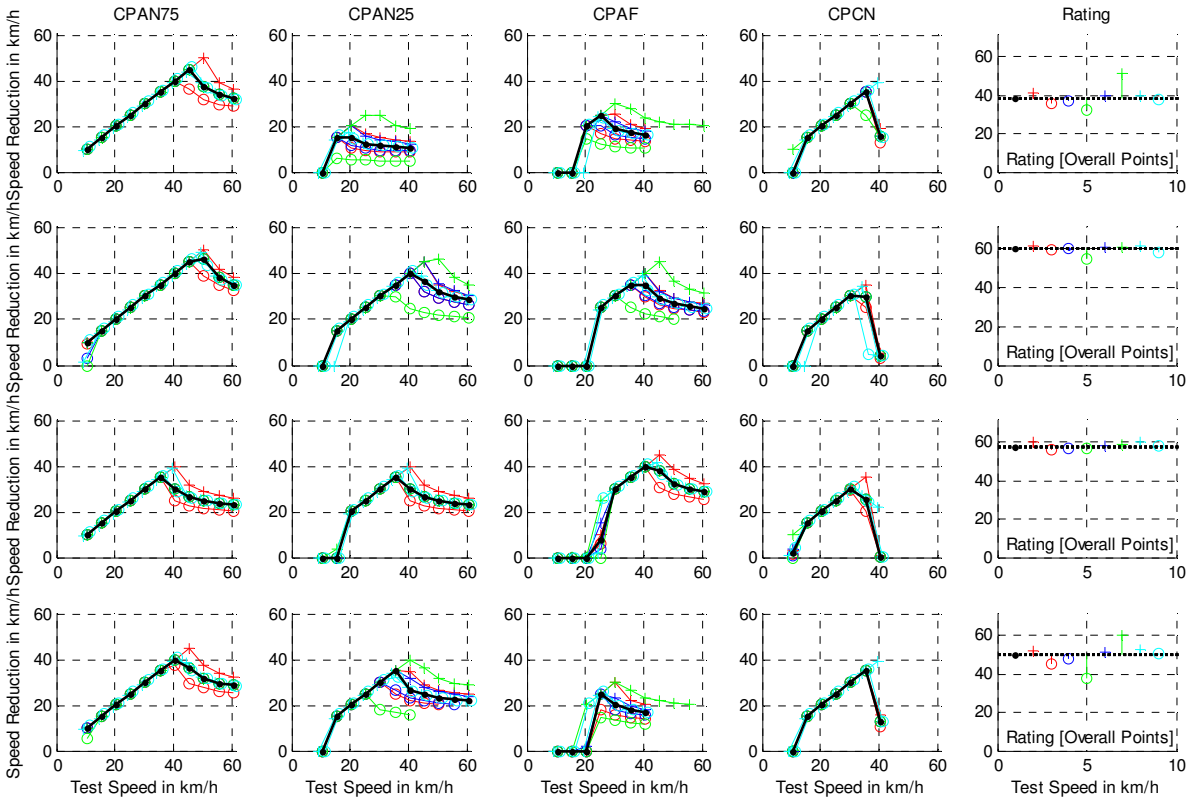


Figure 5. **Results for parameter variations, one vehicle per row. Shown is the achieved speed reduction for each test case from *Table 3* with the appropriate color coding, as well as the overall test result using the Euro NCAP assessment method.**

These results show that the most relevant parameter with respect to test execution precision seems to be the variation in impact point location. While a shift in impact point of 5 cm (as required according to current Euro NCAP protocol) changes the overall results by maximum +4% and -5%, a shift by 20 cm affects the results by a tremendous +33% or -25%. Speed variations during the test, even up to 1 km/h plus or minus (which is four times the value allowed according to the current Euro NCAP protocol) have an effect comparable to the allowed 5 cm impact point shift. Assumed timing variations due to sensor sampling by ± half of sampling time affect the rating slightly more than the allowed 5 cm impact point shift.
These three parameters have been used for a further detailed study of their influence towards the overall rating result per vehicle, see Figure 6. All three plots are based on the baseline, only the selected parameter has been varied.
General tendencies are as expected:

- Moving the impact point away from the position where the pedestrian comes from generally gives AEB systems more time for detection, classification and braking, thus increasing the overall score. The score is more sensitive towards the impact shift for vehicles with generally lower score.
- Increasing the vehicle speed slightly allows for earlier pedestrian detection in critical scenarios for a given sensor angle, but also slightly takes away potential for speed reduction. Both effects seem to compensate each other in the overall rating result, and this is true for all four vehicles.
- Shifting the sensor detection time within an assumed sample rate to earlier values allows for an earlier target detection and thus braking decision which increases the overall rating. The effect of sensor time shift is different for the various vehicles.

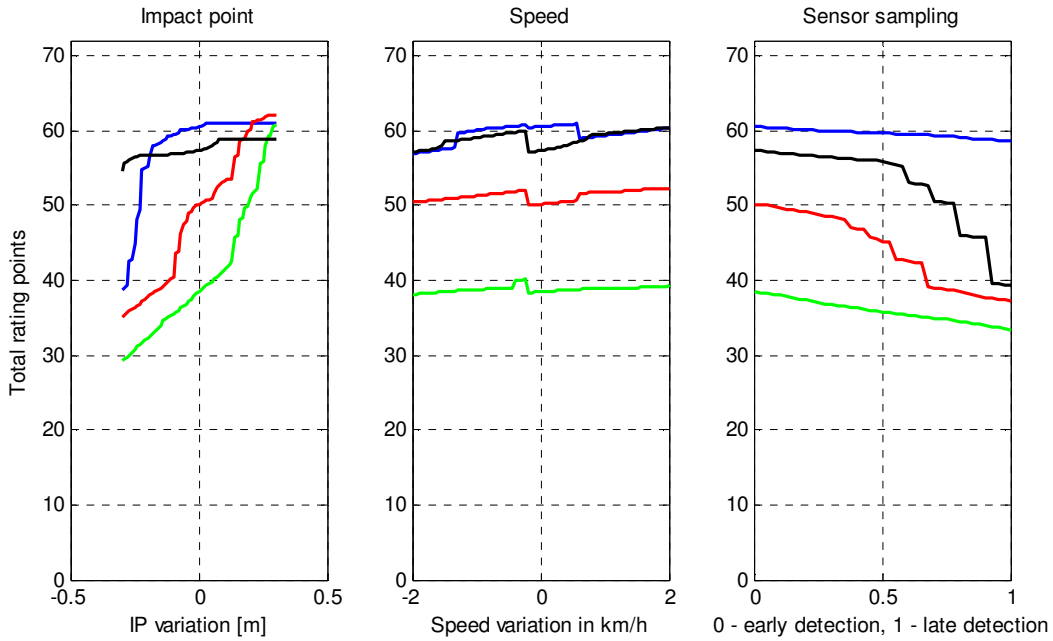


Figure 6. **Sensitivity of overall rating result towards variations of impact point (with respect to the side where the pedestrian comes from), vehicle speed and sensor sampling timing.**

Derivation of required test precision

The most sensitive parameter that needs to be controlled by the test lab is the lateral variation of the impact point. If the test lab is able to maintain the impact point within a range of ± 5 cm, the results in the worst have a variability of ± 2 points of the Euro NCAP rating scheme. For ± 10 cm, this increases to ± 10 points. Note, however, that these values are worst case, assuming a systematic shift of the impact point in the direction of worse test results. In reality, the error will rather be stochastic and vary towards both directions - thus, the influence of the impact point error to overall test result will be much lower.
The other relevant parameter, the vehicle test speed, has a neglectable influence towards the overall score, even if it is varied by ± 2 km/h with respect to the desired test speed.

Consequences

Test experience shows that human drivers are able to maintain one variable (vehicle speed, lateral offset etc.) in acceptable ranges, but are not able to control two variables at the same time. If the test setup is able to synchronize the dummy movement with unprecise but constant vehicle speeds, the tests could be driven by human drivers alone and still allow for an acceptable accuracy with regards to the test result. The influence of vehicle sensor timing issues to the overall test result is then comparable to the error from test conduction.
Since in this case driving robots are not necessary anymore, the test cost can be greatly reduced.

SUMMARY

Goal of this paper was to show a methodology to derive accuracy requirements for a test setup for AEB systems, with focus on AEB pedestrian systems. A simulation model allows to vary relevant parameters and estimate the influence of the given parameters towards speed reduction and overall test result (according to the Euro NCAP assessment method).
The simulation model has been calibrated and validated against AEB pedestrian test data from four different production vehicles.
Sensitivity studies show that the most relevant parameter that needs to be controlled by the test lab is the lateral impact point of the pedestrian dummy on the vehicle-under-test's front. It should be in the range of ± 5 cm. Test vehicle speed variations in the range of ± 2 km/h have an neglectable influence towards test outcome.

Since test vehicle speed just needs to be held constant, but not exactly on a defined value, only one control channel for vehicle control is needed. In this case, it is anticipated that a combination of human driver and test setup (dummy propulsion system) with compensation for speeds and lateral offsets is sufficient to achieve an acceptable test execution precision.

REFERENCES

[AsPeCSS, 2014] Seiniger P, Bartels O, Kunert M, Schaller T: *Test procedure including test target and misuse tests, based on test results*. ASPECSS Deliverable D2.5, May 2014.Available for download at www.aspecss-project.eu.

[Seiniger, 2013]: Seiniger P, Bartels O, Wisch M, Pastor C.: *An Open Simulation Approach to Identify Chances and Limitations for Pedestrian Active Safety*. Paper presented at: 23nd Conference on the Enhancement of the Safety of Vehicles (ESV). May 27-30, 2013; Seoul, South Korea.

[Seiniger, 2015] Seiniger P, Hellmann A, Bartels O, Wisch M, Gail J: *Test Procedures and Results for Pedestrian AEB Systems*. Paper presented at: 24th Conference on the Enhancement of the Safety of Vehicles (ESV). June 8-11, 2015; Göteborg, Sweden.

[Winner, 2011]: Winner H. *Frontalkollisionsschutzsysteme*. In: Winner H, Hakuli S, Wolf G (Ed.).: *Handbuch Fahrerassistenzsysteme*. Wiesbaden: Vieweg + Teubner, 2012:S.522.

Predicting the Impact of Vehicle Safety Developments in Emerging Markets following the Industrialised Countries' Experience

Louise Lloyd, Richard Cuerden, Caroline Wallbank, Matthias Seidl
Transport Research Laboratory, UK

ID#: 15-0239

ABSTRACT

Vehicles in emerging markets are not typically regulated to the same extent as seen in industrialised regions. Casualty rates are considerably higher in these emerging markets, and the lack of vehicle safety regulation is responsible for at least some of the difference. With rapid growth in passenger cars expected, the number of road deaths and casualties in emerging markets is likely to rise, unless targeted and efficient interventions are planned and initiated urgently.

The objective of this work was to quantify the casualty benefit that could be realised in an emerging market, if the experiences and lessons learned in the Europe Union (EU), including minimum car safety standards and consumer testing, were efficiently applied. Malaysia was selected as the emerging market for study in this paper, because it is a Contracting Party to the UN 1958 Agreement and has recently applied the major UN vehicle safety regulations.

Clear differences in vehicle safety developments in Britain compared with Malaysia were identified through analysis of NCAP results and accident data. Frontal impact performance varied, with some cars tested in ASEAN NCAP performing to levels similar to those seen in Euro NCAP today, whilst others were significantly worse. Based on this evidence, in broad terms, this study assumes that new car models sold in Malaysia are approximately 10 years behind today's (2014/15) equivalent European cars, in terms of vehicle safety developments. However, the fleet in Malaysia is older than that seen in Britain, so it is possible that the entire fleet (with many older cars) could reflect a level of safety more like Britain before 2004. Therefore, the Malaysian casualty benefits predicted by this study represent a conservative estimate.

By taking the vehicle safety development experience witnessed in the EU and applying it to the situation in an emerging market, this work quantifies the casualty reduction potential could be a saving of between 1,200 and 4,300 Malaysian fatalities by 2030.

INTRODUCTION

Road vehicles in some of the world's emerging markets are not currently regulated to the same extent as industrialised regions. The combination of growing fleet sizes, made-up of vehicles that do not have to meet basic safety standards, contributes to today's casualty rates, which are considerably higher in emerging markets compared to Europe and other industrialised regions. The lack of vehicle safety regulation is responsible for at least some of the difference. With rapid growth in passenger cars expected to continue, the number of road deaths and casualties in emerging markets is likely to rise, unless targeted and efficient interventions are planned and initiated urgently.

In the automotive sector, minimum safety regulations and standards have evolved in Europe, the US, Japan and other world regions over the last 50 years. There are differences with regard to the nature of tests and criteria which must be met, but the current regulations that apply in the industrialised regions share the objective to provide the highest level of cost-effective safety performance. It is recognised that there are specific differences between the EU and US Regulations, if compared line by line, but it is the outcome mandated by these regulations in real-world car accidents that is important.

Therefore, for practical reasons and to align with the UN Decade for Action for Road Safety, UN Regulations and Global Technical Regulations (GTRs) are referenced as our assumed baseline for minimum mandatory Standards. The UN Decade of Action for Road Safety encourages all countries to apply and promulgate motor vehicle safety regulations as developed by the UN's World Forum for the Harmonization of Vehicle Regulations (WP29). This paper identifies five car regulations, which are defined as a minimum for today's world markets:

- Approved seat belts and anchorages for all seating positions (UN Reg. 14 and 16).
- Occupant protection in frontal collision (UN Regulation 94)
- Occupant protection in side or lateral collisions (UN Regulation 95)
- Pedestrian protection (GTR 9)
- Electronic Stability Control (ESC) (GTR 8)

The UN Decade of Action also has a stated activity to: '*Encourage implementation of new car assessment programmes in all regions of the world in order to increase the availability of consumer information about the*

safety performance of motor vehicles'. The European New Car Assessment Programme (Euro NCAP) was launched at TRL in 1997 to provide an independent assessment of the safety performance of popular cars sold in Europe; the aim was to encourage, on the one hand, consumer awareness of the safety performance of the cars they buy and on the other, to motivate manufacturers to exceed the minimum requirements set out by legislation. Over time Euro NCAP has become more sophisticated, and as well as maintaining its original methodology, which provides a star rating (between one and five stars) based on the car's performance in a number of secondary safety crash tests, including adult and child occupant and pedestrian protection assessments, it also assesses the primary safety or crash avoidance and mitigation credentials of the vehicle.

This paper aims to highlight the combined safety benefits that are realised through the establishment of minimum mandatory car standards enforced through regulation, in conjunction with consumer (NCAP) programmes, which seek to set more challenging vehicle safety design safety targets. Global NCAP has established testing programmes in a number of emerging markets.

Malaysia was selected as the emerging market for this study because it is a Contracting Party to the UN 1958 Agreement and has a vehicle safety improvement strategy, which recently included applying most of the major UN Regulations. Malaysia hosts ASEAN NCAP tests and so it was possible to evaluate the safety performance of some of the vehicles currently sold and to bench-mark performance against similar European vehicles.

Specifically, this paper aims to quantify how many car user fatalities are likely to be prevented in Malaysia between 2014 and 2030, as a combined result of adopting the basic secondary safety measures, namely seat belt standards and UN Regulations 94 and 95 and the impact that ASEAN NCAP will provide on further real world vehicle improvements.

METHOD

Information from Great Britain and Malaysia forms the basis of the study, including accident data, exposure data, and historic and expected developments in fleets. From these data and NCAP test results, the relative safety status of cars in the emerging market compared to the history of European vehicle safety developments is identified. A series of statistical models on British data measure the effect of developments in secondary vehicle safety and apply these effects to the potential impact on casualties in Malaysia if similar developments were seen.

Secondary safety modelling

Car safety has improved in Britain over the past two decades, but how much of this improvement is due to secondary safety features and how much is due to changes in the conditions in which vehicles are driven, for example due to improvements to the road system or changes in the weather, or the way in which they are driven is difficult to define. Statistical modelling is required to disaggregate these effects. The models (Broughton, 2003) use data from the British STATS19 database to model the proportion of casualties killed in road accidents (the severity proportion). The details of the model are shown in the Technical Annex. From these results it is possible to predict the number of casualties which would have occurred if secondary safety had not improved over time.

Prediction modelling

In order to predict the possible future effect of secondary safety improvements in cars in Malaysia we use a scenario based modelling method derived initially in Broughton et al (2000).

The models derived start with a set of possible baseline models, which predict the number of casualties over the next few years if the number of registered vehicles (a proxy for traffic) follows a series of possible scenarios and basic vehicle safety developments continue in the way that has been observed up until the current time. The impacts of further vehicle safety interventions can then be added to the baseline model to predict the impact of these additional measures.

The baseline model is defined as:

$$C'(2030) = C(2012)\ [\ T(2030) / T(2012)\ ]\ (1-\alpha)^{19}$$

where:

- C'(2030) is the predicted adjusted number of car occupant fatalities in 2030 given current progress in vehicle safety
- C(2012) is the number of car occupant fatalities in 2012 (the latest known year)
- T(2030) is the expected number of registered cars in 2030 based on a series of scenarios

- T(2012) is the number of registered cars in 2012
- α is the average annual fatality rate of reduction (adjusted rate) predicted over the 19 year period 2012-2030

For the baseline model, the expected number of registered cars is predicted based on the following broad scenarios:

a) Similar growth or decline to that observed in recent trends.
b) No change in current level.
c) Increased or decreased growth or decline relative to current trend.

Some possible scenarios for changes to vehicle safety progress are added to the baseline model. These new measures have been defined as:

1) Introduction of similar regulations and adaptations to the NCAP testing regime, similar to those seen in Europe over the same time period.
2) Introduction of regulations and NCAP adaptations as above, but over a shorter period.

Application of the scenarios is based on a number of assumptions:

i. Accident types for cars are similar in Britain and Malaysia.
ii. The uptake rate of these regulations and the relative timing is the same in Malaysia as in Britain.
iii. The impact of NCAP in Malaysia is insignificant, but grows as we have seen in Britain in terms of individuals' buying habits. Since the introduction of ASEAN NCAP in Malaysia, we have seen very little evidence of a change in the buying habits in the new sales data as the poorer safety performing cars remain popular.

The results of the secondary safety modelling are used to provide the predicted casualty saving due to secondary saving improvements. In order to apply these to the Malaysian data it is necessary to understand the current state of the Malaysian fleet in terms of fleet renewal and safety standards, relative to a specific point in time in the British fleet. The fleet renewal has been determined from data on the sales of new cars from 2005 to 2010 from Motor Trader (2014). NCAP tests were the most accessible and robust method for comparing safety levels of different cars, however it was not possible to compare star ratings or scores across different NCAPs as the methods vary. Therefore, simple subjective safety levels have been identified by experienced vehicle engineers comparing the visible structural stability of the most popular cars in the most recent NCAP tests in the Malaysian region compared to those in Britain over the last 20 years. These structural scores take into account the visible deformation of the A-pillar and the sill, along with any movement in the steering column and the airbag positioning/ interaction with the dummy's head.

There were some limitations in the method partly due to the video formats. Firstly, these results are based on expert, but subjective opinions. In many cases it was difficult to code exactly into one category and the most appropriate category has been chosen according to the judgment of the experts. Secondly, many of the videos from the different tests in different years and different countries had different views – some just included a side view, others included a frontal view and a view from an internal camera. From some of the camera views it was easier to see steering column and wheel movement compared to others, which may mean that movement was more easily picked up in some types of videos. The NCAP tests are based on the newest car model derivative at the time of the test.

DATA

This section compares vehicle fleet and road fatality figures for Malaysia in the latest available year and for Britain in 1999, 2004, 2009 and 2013. The years covered by the British data reflect the investigation to determine where the current car safety standard in Malaysia is in relation to the British developments.

Casualty data

MIROS provided the total number of road fatalities in Malaysia from 1995 to 2013. This can be combined with information from WHO (2009, 2013) which provide the distributions of these fatalities by road user type for 2007 and 2010. Interpolating from these figures enables us to derive estimates of the number of car occupant fatalities from 2007 to 2012 as shown in Table 1.

Table 1. Estimated number of car occupants from 2007 to 2012 in Malaysia (red numbers are extrapolated)

Year	Deaths	Proportion of car occupants in deaths	Car occupant deaths (estimated)
2007	6,282	23%	1,445
2008	6,527	24%	1,566
2009	6,745	25%	1,686
2010	6,872	26%	1,787
2011	6,877	27%	1,857
2012	6,917	28%	1,937

In 2012, relative to the population the car death rate in Malaysia was 66.1 per million population. In comparison, the highest equivalent rate in Britain was in 1972 (60 per million population) with rates between 31 and 13 from 1999 to 2012.

NCAP comparisons

As of December 2011 new cars in the Asean region could be tested by the New Car Assessment Programme (ASEAN NCAP). A total of nine cars in the top 20 most popular cars in Malaysia were tested by ASEAN NCAP. For the purposes of comparing vehicle safety across Malaysia and Britain, these nine cars and a further 31 cars in the top 10 new sales in 1999, 2004, 2009 and 2013 across Euro NCAP from 1997 to 2014 were observed and scored based on structural deformation in the frontal test which is broadly comparable across the two NCAPs. Structural deformation was classified into one of nine categories:

1. No intrusion
2. Mild A-pillar damage
3. Severe A-pillar damage
4. Mild vehicle sill (base of door) or footwell movement/damage
5. Severe vehicle sill movement/damage
6. Mild A-pillar and sill damage
7. Severe A-pillar damage and mild sill damage
8. Mild A-pillar damage and severe sill damage
9. Severe A-pillar and sill damage

The distribution of the structural deformation score is shown in Table 2. This shows a mixture of scores across the different NCAP regions and years.

Table 2. Count of cars tested by NCAP by structural deformation score

Structural deformation	1	2	3	4	6	7	9	Proportion of new sales
Asean	4	3	1				1	45%
Euro 1999		1	2			3	4	35%
Euro 2004	4	1	1		1	1		26%
Euro 2009	3			2				16%
Euro 2013	5	1		2				24%

Broadly, the Asean cars appear to be performing somewhere around the Euro 2004 level.

Fleet data

Table 3 compares car fleet facts in Malaysia in 2012 and Britain in 2004. In Malaysia, 10.5m cars were registered which included 0.6m new cars (OICA (2013) for registrations and Malaysian Automotive Association (MAA, 2014) for sales data). The number of registered cars is growing substantially each year.

Table 3. Comparison of registered car facts in Malaysia and Britain

	Britain: 2004	Malaysia: 2012
No. reg cars (DfT, 2013a)	25.8m	10.5m
Growth in reg cars from previous year	3.1%	6.5%
Proportion of reg cars new (DfT, 2013b)	9.7%	5.3%
Proportion of cars aged 6 years or less (DfT, 2014)	53%	28% (2010)
Motorisation rate (DfT, 2013b)	0.44	0.36

Detailed registered car data were not available to the research team for Malaysia during the research. However, based on sales data from 2005-2010 (Best Selling Cars Blog, 2014) we estimated that 28% of the Malaysian car fleet is at most six years old. In comparison, the number of cars and the proportion of these cars that are 6 years or younger is twice as high in Britain in 2004. Given the information in Table 3, there is some suggestion that the current fleet in Malaysia contains older cars relative to Britain in 2004, therefore any expected casualty savings are likely to be underestimates.

RESULTS

Secondary safety in Britain

The secondary safety modelling is shown in detail in the Technical Annex. This shows that there have been considerable reductions in fatalities in Britain since 1990 due to developments in secondary safety. It is estimated based on these models that 302 lives have been saved between 2004 and 2013 in Britain due to secondary safety which is a reduction in car driver fatalities of 4.2%.

Predictive modelling in Malaysia

Based on the number of registered cars and the car occupant fatality estimates shown in Table 1, the fatality rate for car occupants can be estimated. These are shown in Figure 1 with two possible trends. Both trends are based on a linear function. A longer term analysis (1995 – 2013) was carried out on all fatalities and all registered vehicles and this suggested that a linear trend was more appropriate than a log linear trend. There is a longer trend which takes into account data from 2007 to 2012 and a shorter trend (based on 2009-2012), which encompasses what appears to be a possible change in trend in 2009, when the car occupant fatality rate started to decrease. It was decided that, given these trends are based on relatively few data points, the longer trend should be used as this is less likely to contain as much random fluctuation as the shorter trend.

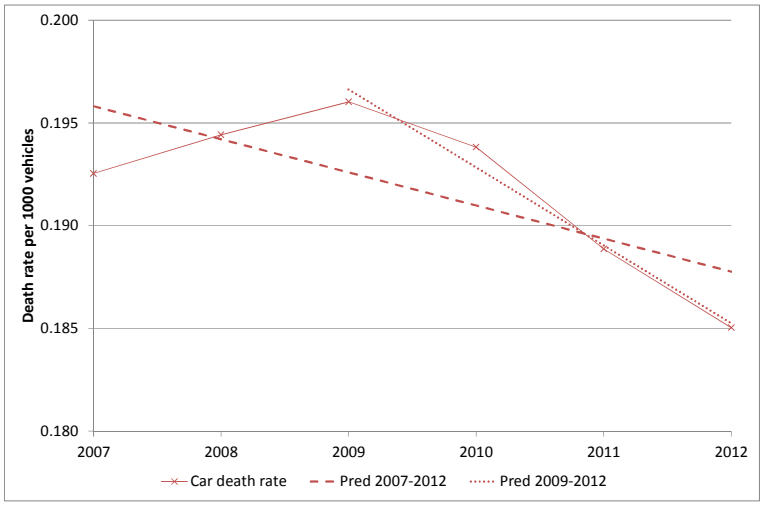


Figure 1. Car occupant fatality rate per 1000 registered cars in Malaysia (2007-2012)

The possible baseline scenarios for car registration growth in Malaysia from 2012 to 2030 have been devised as:

A. The trend in car registrations continues to grow linearly at the current rate: an average annual rate of 5.4% relative to 2012
B. The trend in car registrations continues to grow linearly at an average annual rate of 0.7% relative to 2012. This is equivalent to the growth rate in Britain between 2004 and 2013.
C. The trend in car registrations continues to grow linearly at an average annual rate of 14.8% relative to 2012. This is equivalent to the highest annual average increase in car registrations in Britain, in the 1950s.
D. The trend in car registrations continues to grow linearly at an average annual rate of 5.4% as in scenario a) but also encompasses a gradual move from motorcycles to cars resulting in half of the number of motorcycles in the fleet in 2030. This is equivalent to an annual average increase of 7.0% per year.

Relative to the population (and assuming a constant linear growth), the associated motorisation rates by 2030 are 0.52 (A), 0.29 (B), 0.96 (C) and 0.59 (D). The current motorisation rate in Malaysia is 0.36 and in Britain is 0.45

as shown in Table 3. This shows that scenario C is highly unlikely, and further calculations do not consider this scenario.

Assuming that the relationship between car occupant fatalities and registered cars remains the same as in the current trend, the derived number of car occupant fatalities that could be expected under the three feasible scenarios A, B and D are shown in Figure 2. The associated increase in car occupant fatalities in 2030 compared to 2012 for the three remaining baseline scenarios 68% (A), -5% (B) and 92% (D).

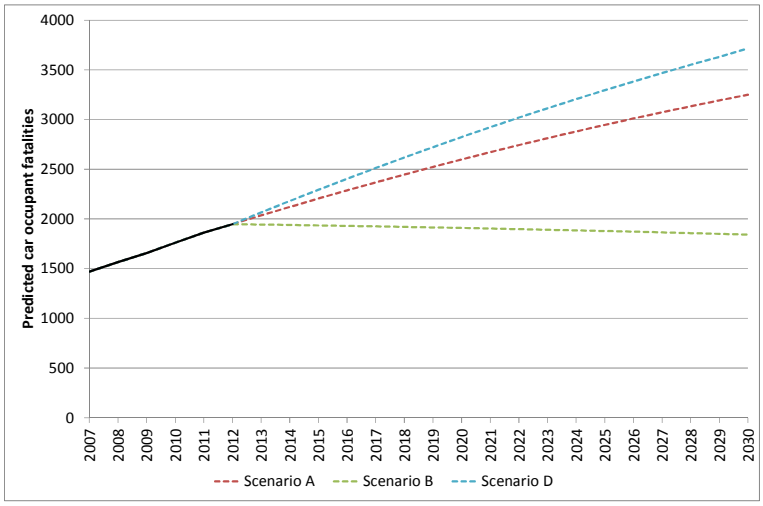


Figure 2. Predicted number of car occupant fatalities in Malaysia 2007-2030 based on three scenarios

These baseline scenarios assume that vehicle safety developments continue at the current rate. However, if the impact of regulations on vehicle manufacturers and NCAP on buying habits that were seen in Europe could be replicated in Malaysia, then we would expect a reduced number of fatalities over the same period. Based on the comparisons shown in vehicle safety between Malaysia and Britain on new cars tested in NCAP programmes, the suggestion is that the current vehicle safety standard in Malaysia could be equivalent to the 2004 standards in Europe. We know that Malaysia is not exactly as Europe was in 2004. In particular, as shown in Table 3, the proportion of the cars in the fleet that were reasonably new (6 years or under) is approximately half that in Britain in 2004. This suggests that the fleet renewal rate in Malaysia is currently slower than observed in Britain in 2004, and therefore any regulations or new safety technologies or features would likely infiltrate the fleet at a much slower pace than observed in Britain. As discussed above, in Britain, 4.2% fewer car drivers died between 2005 and 2013 due to the progression in secondary safety.

We have predicted the impact on fatalities of two future scenarios:

1. A proportional reduction in the baseline scenarios fatalities equivalent to half the proportional reduction observed in Britain to take into account the difference in fleet turnover rate observed in Britain.
2. The same proportional reduction in the baseline scenarios as observed in the Britain which might reflect a quicker fleet turnover or a quicker uptake of regulations than observed in Britain.

Table 4 shows the estimated potential car occupant fatality savings over 17 years between 2014 and 2030 based on the potential baseline trends in car occupant fatalities in Malaysia shown in Figure 2, and the two named future scenarios 1 and 2 above.

Table 4. Estimated potential car occupant fatality savings between 2014 and 2030 given future scenarios 1 and 2

	Future scenario 1		Future scenario 2	
Baseline scenario	**Potential total fatality saving**	**Potential fatality saving proportion**	**Potential total fatality saving**	**Potential fatality saving proportion**
A	1,914	4.1%	3,827	8.3%
B	1,225	3.8%	2,449	7.6%
D	2,139	4.2%	4,278	8.4%

CONCLUSIONS

In 2012, the car user death rate was 66.1 per million Malaysian population compared to 13 per million British population. This represents a significantly higher casualty rate, and this is only partially associated with differences in vehicle safety standards.

Comparing the performance of cars in Asean and Euro NCAP frontal impact tests, we concluded that broadly, today's Asean cars could be performing somewhere around the Euro 2004 level. This was based on an engineering visual assessment of the cars' structural behaviours during the impact tests. However, it is not known how well the Asean cars would perform in equivalent Euro NCAP side impact conditions, or the level of secondary safety offered by cars in the current fleet, which have not undergone NCAP testing.

The casualty modelling assumes that recent car safety regulations that have come into force in Malaysia, have improved the secondary safety protection afforded to new car model users, to a level similar to that seen in Europe in 2004. Based on information on the age of the car fleet in Malaysia, it is possible that the entire fleet (with many older cars) could reflect a level of safety more like Britain before 2004. Therefore, the casualty benefits predicted for Malaysia by this study represent a conservative estimate.

In Britain from 2005-2013, the estimated effectiveness of improvements in car secondary safety, since 2004-05 registered vehicles, was 4.2% for car driver fatalities. That is, 4.2% fewer car drivers died between 2005 and 2013 due to improvements in secondary safety in cars.

The current motorisation rate in Malaysia is 0.36 and in Britain is 0.45. Depending upon the change in motorisation rates and the uptake of new cars, we estimate that between 1,200 (scenario B, 0.29 motorisation rate) and 4,300 (scenario D, 0.59 motorisation rate) car user fatalities could be prevented on Malaysia's roads between 2014 and 2030. This assumes that the car secondary safety improvements seen in Britain since 2004 are transferred to the Malaysian vehicle fleet.

This paper has provided estimates for fatality reductions associated with secondary safety measures for car users that are likely to be seen in Malaysia because of the adoption of UN Regulations. Establishing minimum mandatory regulations within the marketplace has been a significant step and provides a real world example of what can be achieved.

Further work will summarise the potential benefits for pedestrian protection and ESC advances in Malaysia and other emerging markets.

Acknowledgements
This research was funded by Global NCAP.

References

Broughton, J. (2003). *The benefits of improved car secondary safety.* Accident Analysis & Prevention, Volume 35, Issue 4, 527-535.

Broughton J, Allsop RE, Lynam DA and McMahon CM (2000). *The numerical context for setting national casualty reduction targets.* TRL Report 382, Transport Research Laboratory, Crowthorne.

Broughton J and Keigan M (2010). *Linking STATS19 and Scottish Hospital In-Patient data for the SafetyNet project.* TRL Report PPR207, Transport Research Laboratory, Wokingham.

Department for Transport (DfT) (2013a). *Vehicles (TSGB09) - Statistical data sets - GOV.UK. [ONLINE]* Available at: https://www.gov.uk/government/statistical-data-sets/tsgb09-vehicles#ASEAN NCAP-TSGB0901-VEH0103 [Accessed 4 Mar. 2015].

DfT (2013b). *Vehicles (TSGB09) - Statistical data sets - GOV.UK. [ONLINE]* Available at: https://www.gov.uk/government/statistical-data-sets/tsgb09-vehicles#ASEAN NCAP-TSGB0901-VEH0103 [Accessed 4 Mar. 2015].

DfT (2014). *Number of registered cars by make, model and year of first registration.* Data direct from DfT.

Gasnier, M (2014). *Best Selling Cars – Matt's blog » 2014. [ONLINE]* Available at: http://bestsellingcarsblog.com/2014 [Accessed 4 Mar. 2015].

Lloyd, L.K., Reeves, C., Broughton, J. and Scoons, J. (2012). *Investigating the reduction in fatal accidents in Great Britain from 2007-2010.* TRL Published Project Report PPR663.

Motor Trader (2014). *New Vehicle Sales by brand. [ONLINE]* Available at: http://www.motortrader.com.my/ . [Accessed 02 March 15].

OICA (2013). *PC World Vehicles In Use. [ONLINE]* Available at: http://www.oica.net/wp-content/uploads/2013/09/pc-wviu.pdf . [Accessed 03 March 15].

MAA (2014). *Malaysia Automotive Info – Summary of sales and production data. [ONLINE]* Available at: http://www.maa.org.my/info_summary.htm . [Accessed 03 March 15].

WHO (2009). *Global status report on road safety 2009. [ONLINE]* Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2009/en/ [Accessed 4 Mar. 2015].

WHO (2013). *Global status report on road safety 2013*. [online] Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/ [Accessed 4 Mar. 2015].

TECHNICAL ANNEX – Secondary safety modelling

Method

The following variables are included in logistic regression models:

- **Registration year of the vehicle**: used to estimate the reduction in the severity of drivers' injuries linked to changes in succeeding 'cohorts' in the car fleet
- **Accident year**: accounts for the fact that other road safety measures and conditions will have affected the road system
- **Age and sex of the casualty**: to allow for the differences in injury severity and car choice for different demographics of drivers
- **Vehicle type of striking car**: controls for the differences in protection offered to occupants of different sizes of cars.
- **Road type**: is a proxy for speed of accident and controls for the influence of speed on injury severity.

The modelling is used to predict the number of casualties which would have occurred if secondary safety had not improved. This calculation assumes that improving car secondary safety cannot prevent occupants from being injured in an accident, but can reduce the severity of the injuries suffered. As a result, this model is likely to underestimate the actual benefit as some car occupants who would previously have been slightly injured may now not be injured.

Results

In total, 2,802,648 car driver casualties were injured in accidents recorded in STATS19 between 1989 and 2013. Of these 83% (2,329,292) had a valid age, vehicle registration year, vehicle type and speed limit and these data have been used for this analysis. It has been assumed that excluding those vehicles where the data is unknown does not bias the results of the model, and the casualty estimates presented are weighted to take into account those that are unknown.

Modelling results Figure 3 shows the proportion of male 25-59 year old small family car driver casualties killed by registration year in 2013 on built up and non-built up roads. The blue dots show the result of the model when year of registration is included as a factor, and the red line shows the model results when a linear trend is assumed for the improvements to secondary safety. A break-point in the linear trend occurred in 1990-91 registered vehicles on both built-up and non-built-up roads.

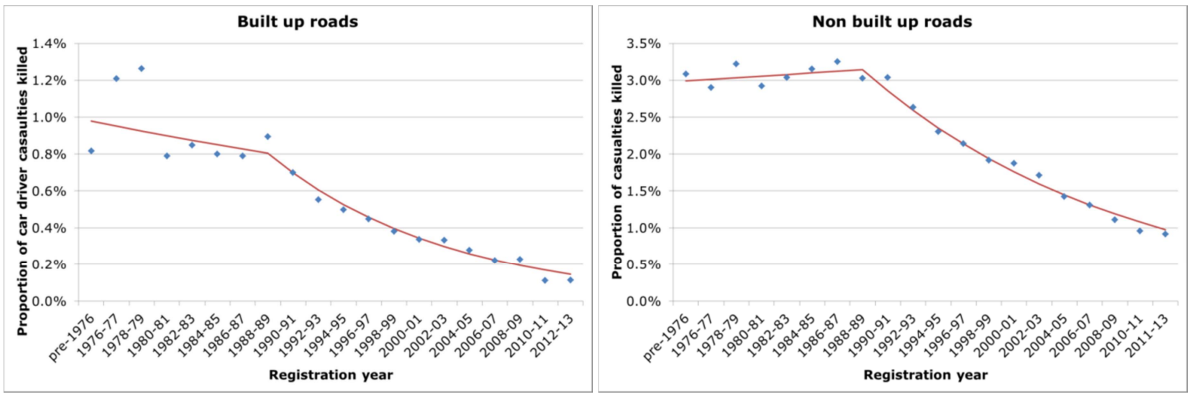


Figure 3: Modelled fatality proportion by registration year for car drivers (males 25-59, small family cars, 2013)

Figure 3 indicates that, for accidents on built-up and non-built-up roads, since 1990-91, improvements to secondary safety for car occupants have improved more rapidly than pre-1990. It is suggested that this is due to the impact of manufacturers starting to consider the impact of possible regulations on the car industry.

One other possible explanation for this trend is a change to the level of underreporting in STATS19. If slight and/or serious casualties were being reported less over time, but the level of fatality reporting remained the same then the proportion of casualties killed would appear to increase. In June 2006, the Department for Transport commissioned some work to match STAT19 data to hospital data in an attempt to determine levels of underreporting (Broughton and Keigan, 2010); however the conclusions are hard to draw out and the trend over time was not investigated. Even if this trend is true, the changes in reporting rate do not bias the crucial results relating to registration year as we are interested in the trend from 2004 onwards when applying the data to Malaysia.

Estimated casualty benefits Following the methodology outlined in Broughton (2003), the models described above can be used to estimate the number of lives that have been saved by improvements to the secondary safety of cars. For example, it is possible to determine if the safety of cars had remained at the level of the 2004-05 registered cars how many additional fatalities would have occurred up to 2013.

These calculations assume that the total number of collisions remains unchanged, but that more drivers would have been killed because of the lower secondary safety. The model is used to adjust the severity proportions of the modern cars to match with those registered in 2004-05. Casualties in cars registered before 2004-05 are assumed to remain unaffected.

Improvements to secondary safety are likely to have reduced the total number of casualties as some casualties who would have previously been slightly injured in the collision are not injured in more modern cars. As a result, the casualty estimates presented in Table 5 are an underestimate for the actual casualty benefit. The total figure shown in Table 5 is based on modelled data to smooth out variability in the trend due to random variation.

Table 5: Actual and estimated car driver fatality numbers in 2007, 2010, 2013 if secondary safety had remained at level of 2004-05 registered vehicles

Accident year	Actual fatality numbers	Estimated fatality numbers if secondary safety had not improved	Reduction in fatalities due to secondary safety improvements	Proportional reduction due to secondary safety
2007	934	990	56	5.7%
2010	568	583	15	2.5%
2013	543	602	59	9.8%
Total (modelled) 2005-2013	6,874	7,176	302	4.2%

Hence, between 2005 and 2013 the estimated effectiveness of improvements in secondary safety since 2004-05 registered vehicles was 4.2% for car driver fatalities. That is, 4.2% fewer car drivers died between 2005 and 2013 due to improvements in secondary safety in cars.

CONSUMER SAFETY INFORMATION PROGRAMS AT IIHS

David Zuby
Insurance Institute for Highway Safety
United States

Paper Number 15-0228

ABSTRACT

Since 1969, when the Insurance Institute for Highway Safety (IIHS) began publishing results of low-speed crash tests to highlight differences in vehicle bumpers, it has been a significant source of information about how the safety of different vehicle designs varies. Currently, IIHS maintains crashworthiness ratings covering five crash modes along with ratings of front crash prevention (FCP) systems and children's booster seats, as well as annual updates of insurance loss reports from its affiliate, the Highway Loss Data Institute (HLDI).

This report describes the experience with IIHS's latest consumer information efforts and identifies the next areas of consumer information to come online. It presents information about the number of vehicle models and booster seats evaluated; their ratings assigned as well as media, consumer, and manufacturer response; and small overlap crashworthiness and FCP ratings. Research underpinning future rating programs addressing Lower Anchors and Tethers for Children (LATCH) and advanced head lighting systems also is summarized.

Since launching its booster seat ratings, IIHS has evaluated 200 designs for their ability to adjust rear seat belt fit to booster-age children across a wide variety of rear seat belt configurations. The number of models rated Best Bet, indicating they will provide good belt fit in common passenger vehicles, has increased from a low of 10 in 2008 to 69 in 2014. Media coverage of these annual ratings announcements is estimated to average an audience of 88 million people in the United States. IIHS internet pages with booster ratings are among the most viewed, with an average of 102,800 page views monthly.

IIHS began rating vehicle front crashworthiness on the basis of a 64 km/h small overlap crash against a rigid barrier in 2012. Of the 118 currently rated 2015 models, 49 are good, 25 acceptable, 23 marginal, and 21 poor. Several models have been tested in two design iterations with improved performance in the second test, indicating automakers are able to design vehicles to better protect occupants in similar crashes. It is estimated that the media coverage across all small overlap ratings announcements has achieved 1.1 billion views. Surveys of automobile dealers indicate that good ratings in this test have led to increased sales, at least in the short term.

IIHS ratings of vehicle FCP systems include both warning and autobraking functions. The proportion of new models available with FCP of any kind has increased from 30 to 60 percent. The combined media coverage of three announcements featuring FCP ratings were viewed 212 million times. While not as strong as for crash test ratings, there was indication that these announcements positively affected sales of vehicles with these systems.

Large audiences for IIHS consumer information programs have prompted manufacturers of rated products to make changes in ways indicated by IIHS tests. Based on this experience with current programs, there is good reason to believe that IIHS ratings of LATCH and advanced head lighting systems can also improve vehicle safety.

INTRODUCTION

The Insurance Institute for Highway Safety (IIHS) was founded in 1959 as a nonprofit research and communications organization. Its mission is to conduct and publish research that will lead to the reduction of deaths, injuries, and property damage associated with crashes on roads in the United States of America. One of the ways in which IIHS uses its research to motivate improvements in vehicle designs has been through consumer information programs that

highlight differences in safety among different vehicles. The most recognizable form of such programs are crashworthiness ratings based on crash tests, the first of which were published by the New Car Assessment Program of the U.S. Department of Transportation [1]. However, such consumer information can take other forms such as the publication of insurance losses by make and model, which have been produced by IIHS affiliate the Highway Loss Data Institute (HLDI) since 1973 [2]. More recently, IIHS has applied the consumer information model to motivate improvements to child restraint systems and commercial truck underride guards.

IIHS's first foray into providing information about crash differences among contemporary vehicle models was bumper tests in 1969, which ultimately led to regulatory bumper performance requirements. In 1972, HLDI was formed to examine highway safety issues through analysis of insurance loss data provided by a subset of IIHS sponsors. The compilation of data from the largest insurers in the United States allows HLDI analyses to ascertain the model-specific contribution to losses under different types of insurance coverage. HLDI has published annual comparisons of losses by make and model since 1973. While not fully explaining the differences in insurance costs paid by different insurance customers, these analyses partly explain the differences an individual will pay depending on the model insured. Results are published on a scale that relates the insurance experience of a specific model to the average for all contemporary passenger vehicles. This type of information was considered so important to consumers that the Motor Vehicle Information and Cost Savings Act of 1972 required the U.S. Department of Transportation to devise a means of distributing it wherever new vehicles were sold [3]. From 1993 until 2013, when the availability of information via the internet obviated the need for auto dealers to keep printed brochures on hand, HLDI's data have helped the National Highway Traffic Safety Administration fulfill this requirement [4].

Since 1995, IIHS has created a rigorous program of crash test ratings that have led to measureable improvements in the crash protection offered by modern vehicles. Its first crash rating program involved front crashes at 64 km/h with a 40 percent overlap against a deformable barrier. Test results are graded on a scale of good, acceptable, marginal, and poor to reflect the relative protection for occupants exposed to similar crashes. The ratings are based not only on measurements made by sensors in the test dummy but also on analysis of the dummy's observable motion and measurements of safety cage deformation. Figure 1 illustrates the increasing availability of good-rated vehicles across model years, and these improvements have been associated with a lower risk of death in front crashes [5].

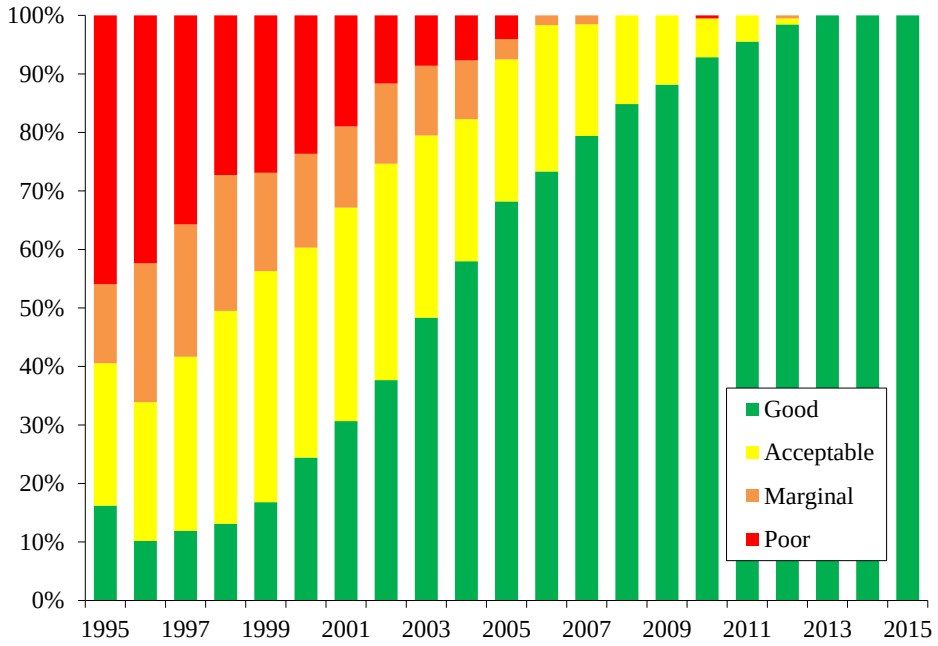


Figure 1. IIHS moderate front overlap crash test ratings by model year.

The IIHS crashworthiness program currently consists of five tests, all using the poor-to-good scale of ratings: two front (moderate and small overlap impacts at 64 km/h), a side impact at 50 km/h, a quasi-static test of roof strength, and a simulated rear crash used to evaluate the ability of vehicle seats and head restraints to mitigate neck injury. Like IIHS's original front crash test, the side and rear crashworthiness evaluations have resulted in measurable changes in the likelihood that vehicle occupants will be injured or killed in crashes represented by the corresponding evaluation [6][7]. The roof strength test was initiated based on research establishing the protective effect of strong roofs in rollover crashes [8][9]. All of these programs have led to meaningful improvements in the safety of the U.S. passenger vehicle fleet.

IIHS *Top Safety Pick* awards were launched in 2005 to recognize those 2006 model year designs that offered the highest levels of crash protection, as indicated in IIHS tests. In addition to rewarding automakers whose products offer the highest level of safety, the rewards also generate a single list of recommendations to consumers. The annual announcement of these ratings is widely covered by the U.S. news media. Table 1 shows the number of times an IIHS-produced video was broadcast for each of the *Top Safety Pick* announcements since 2006, along with the estimated size of the audience that saw those broadcasts. Also, automakers increasingly use *Top Safety Pick* claims in their advertising, as Table 2 attests. Public attention to this award amplifies the incentive for automakers to improve the safety of their products.

Table 1.
Number of television broadcasts and estimated audience for IIHS *Top Safety Pick* announcements.

Award year	Number of broadcasts	Estimated audience (millions)
2006	1,123	95.0
2007	1,365	102.3
2008	1,051	100.0
2009	1,880	115.1
2010	2,580	149.4
2011	2,052	114.2
2012	637	41.3
2013	1,262	155.5
2014	856	74.8
2015	1,083	76.7

Table 2.
Number of requests to approve *Top Safety Pick* advertising claims.

Award year	Advertising approval requests
2006	218
2007	288
2008	361
2009	328
2010	341
2011	646
2012	635
2013	573
2014	614

The criteria for earning *Top Safety Pick* was made more stringent twice between its introduction and 2012. The *Top Safety Pick+* was introduced in 2014 as a type of "soft landing" for the introduction of new front crashworthiness requirements. Designs that met the previous year's criteria kept their *Top Safety Pick* accolades, but an acceptable or good rating in the new test was required to earn the new award. Table 3 shows the criteria for *Top Safety Pick* and *Top Safety Pick+* awards during their history. For the foreseeable future, *Top Safety Pick+* will continue to be the way to identify IIHS's latest safety recommendations.

Ratings of small overlap front crashworthiness and front crash prevention systems are the latest additions to IIHS consumer information programs encompassed by *Top Safety Pick.* Also, IIHS has rated belt-positioning booster seats for children since 2009 and has published a one-time comparison of commercial truck underride guards. Evaluations for vehicle LATCH (Lower Anchors and Tethers for Children) and advanced headlights are also being developed. The remainder of this paper describes these programs and the safety improvements they have compelled or intend to once launched.

Table 3.
IIHS *Top Safety Pick* criteria, 2006-15.

Award year	Moderate overlap front crash	Side impact	Rear crash	Electronic stability control	Roof strength	Small overlap front crash	Front crash prevention
2006	Good	Good	Good or acceptable				
2007	Good	Good	Good	Available			
2008	Good	Good	Good	Available			
2009	Good	Good	Good	Available			
2010	Good	Good	Good	Available	Good		
2011	Good	Good	Good	Available	Good		
2012	Good	Good	Good	*	Good		
2013	Good	Good	Good		Good	Acceptable	
TSP+	Good	Good	Good		Good	Good	
2014	Good	Good	Good		Good	Good or acceptable	
TSP+	Good	Good	Good		Good	Good or acceptable	Basic or better
2015	Good	Good	Good		Good	Good or acceptable	
TSP+	Good	Good	Good		Good	Good or acceptable	Advanced or better

TSP+ = Top Safety Pick+
*Electronic stability control became required by regulation beginning September 1, 2011

LATEST CONSUMER INFORMATION PROGRAMS

Booster Seat Belt Fit Ratings

Booster seats are intended to improve the fit of vehicle seat belts for children who have outgrown child restraint systems with internal harnesses and that are anchored to the vehicle with the seat belts or dedicated attachments. The simplest boosters are cushions that raise the child above the vehicle seating surface such that the lap belt crosses the bony pelvis and the shoulder belt does not interfere with the neck or face. This improves the comfort for young vehicle occupants and, more important, helps the seat belts restrain them more effectively in a crash. Children riding in boosters are much less likely to be injured in a crash than children secured by vehicle seat belts alone [10].

Due to the variation in rear seating areas of modern vehicles and the variety of booster seat configurations, some combinations do not result in good belt fit for some children. This complication leads to a high rate of booster misuse in the field [11]. To address this problem, IIHS researchers worked with the University of Michigan Transportation Research Institute (UMTRI) to develop a method for evaluating which booster seat designs provide the best belt fit across a range of vehicles [12]. The evaluation uses a 6-year-old Hybrid III dummy to represent the relevant child population. The dummy is seated on each booster, which is installed on a test fixture representing a rear seat with variable seat belt anchor positions. Both lap and shoulder belt positions on the dummy are recorded for four seat belt anchor configurations representing the variation in the modern vehicle fleet. Boosters that provide ideal fit — lap belt lying flat on the thighs and close to the hip with the shoulder belt crossing the middle of the sternum — in all four configurations are rated Best Bet, reflecting the notion that the design provides sufficient control of the belt routing that it would be expected to provide correct fit for all booster-age children in almost any vehicle. Boosters rated Good Bet also would be expected to provide no more than slight deviations from ideal fit in almost any vehicle. The Check Fit rating means that acceptable fit is possible for some, but not all, seat belt anchor configurations. Consumers are advised to check whether such boosters provide good belt routing for their children in the vehicles in which the boosters will be used. Seats rated Not Recommended fail to provide acceptable belt fit in any of the four tested belt anchor configurations.

The booster seat belt fit evaluations were first published in 2008 and have been repeated every year since. Among the 41 booster designs evaluated in the first year, 10 were rated Best Bet, 5 were Good Bet, and 13 were Not

Recommended. The rest were rated Check Fit. The availability of Best Bet boosters has increased in every update of the IIHS ratings, indicating that child restraint manufacturers are taking the rating guidance into account when developing new products. Figure 2 compares the current distribution of seat ratings with those from 2008. These improvements also have led to customer satisfaction for those boosters earning Best Bet ratings, as this email correspondence from Michael Noah, president of Harmony Juvenile Group, attests:

> I'm happy to report that over the past year many customers with some of our Best Bet boosters experienced only good outcomes from some pretty severe crashes. It is nice to know that the combination of technology in the seat design along with the good work that you do — truly leads to the outcomes we are all working towards.

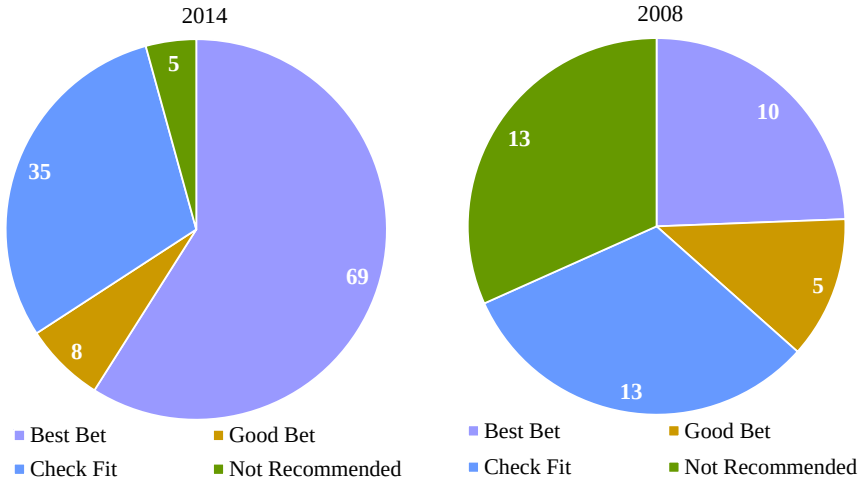


Figure 2. IIHS Booster seat ratings published in 2014 compared with those published in 2008.

The public is very much interested in information about child safety. The booster ratings pages of the IIHS internet website are among the most frequently visited, receiving an average of 102,800 page views monthly. Likewise, the television audiences for announcements of new booster ratings are regularly large. Table 4 shows the number of television broadcasts and estimated audience for each update of IIHS booster ratings.

Table 4.
Number of television broadcasts and estimated audience for IIHS booster seat ratings announcements.

Award year	Number of broadcasts	Estimated audience (millions)
2008	1,384	87.1
2009	3,669	182.9
2010	2,575	153.3
2011	739	44.5
2012	864	42.6
2013	1,452	53.7
2014	1,299	53

Truck Underride Guards

During the past 5 years, the number of passenger vehicle occupants killed each year in crashes with large trucks has averaged more the 2,400 in the United States. According to a 1997 study, about half of these deaths involve the passenger vehicle underriding some portion of the truck [13]. Approximately one-fifth of the underride deaths occur in crashes to the rear of the truck, despite requirements that many heavy trucks be equipped with underride prevention guards.

Research by IIHS identified the nature of underride guard failures, and testing showed there was a range of effectiveness among designs meeting U.S. regulatory requirements [14]. Specifically, two of the three tested guards prevented underride of a midsize car in full-overlap crash tests at 56 km/h, and only one prevented underride when overlap with the guard was reduced to 50 percent [15]. This research received moderate amount of media attention and was featured in 1,296 broadcasts with a total estimated audience of 64 million viewers.

Follow-up testing during 2012-13 involving truck trailers from the eight largest manufacturers in the U.S. market already showed improvements for the guards that performed least well in earlier testing [16]. All eight guards prevented underride in the full-overlap test, all but one prevented underride in the 50 percent overlap condition, and only one prevented underride when the overlap was reduced to 30 percent. This test program also received media attention and was featured in 811 broadcasts with a total estimated audience of 37 million viewers. This attention has resulted in six of the seven trailer manufacturers whose guards could not prevent underride in all three test conditions to report they are developing more effective guards. In 2014, one company, Vanguard, whose guard failed to prevent underride in the 50 percent overlap condition in both rounds of tests, tested an improved design at IIHS that successfully prevented underride in this condition. These observations suggest that comparative safety evaluations can be an effective means of stimulating improvements in the commercial fleets as it has for private passenger vehicles.

Small Overlap Front Crashworthiness Ratings

By the 2008 model year, 85 percent of passenger vehicles rated by IIHS earned good ratings in the moderate overlap front crash test and the remainder were rated acceptable, indicating the evaluation was no longer providing significant discriminating information to consumers nor spurring further improvements in front crashworthiness design. So IIHS researchers examined crashes of good-rated vehicles to ascertain whether further improvements in front crashworthiness were possible. This research focused on crashes resulting in serious injury or fatality and found that many had damage patterns indicating overlaps smaller than the 40 percent overlap of the crash tests [17]. In such crashes, the energy absorbing structures in the crush zones often were minimally damaged while the safety cages were severely collapsed. Crash testing showed that some vehicles provided better protection with less collapse of the safety cage than others in the same small overlap test configuration. This led to the development of IIHS's newest crash test evaluation, which involves crashing a vehicle at 64 km/h with 25 percent of its width overlapping a rigid barrier with a rounded edge (Figure 3).

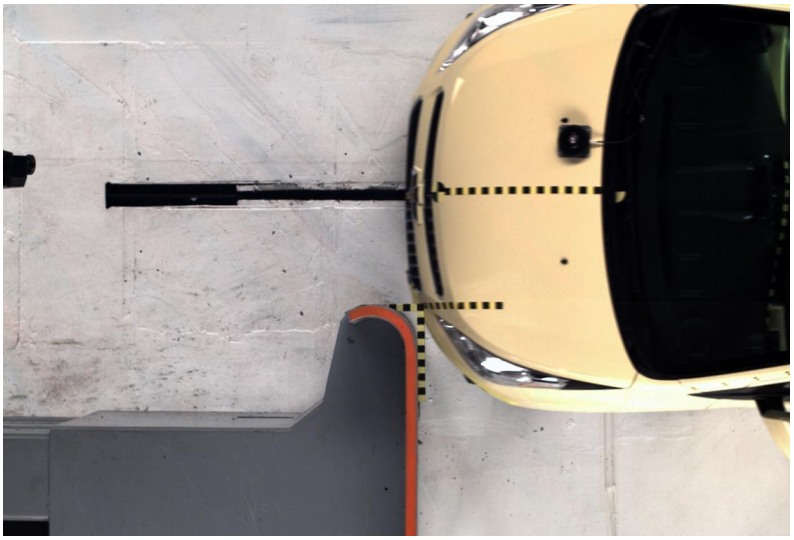


Figure 3. Small overlap crash test configuration (overhead view at t = 0).

Small overlap front crash test ratings of new vehicles were first published in August 2012. Eleven midsize luxury cars were tested, with two rated good, one acceptable, four marginal, and four poor. Since that time, 133 model designs have been evaluated in this test. Figure 4 shows an example the survival space difference for a vehicle rated good compared with one rated poor. In addition to structural differences, these tests also show differences in the ability of restraint systems to prevent impacts between the driver dummy's head and instrument panel (Figure 5). Figure 6 shows ratings in the small overlap front test by vehicle type for the 2015 model year. Eight models have been tested both before and after modifications or redesigns intended to improve front crashworthiness. All eight modifications/redesigns showed increased survival space compared with their predecessors, and six of the eight had improved overall ratings [18]. It is too early to know the extent to which such changes are affecting real crash outcomes.

These new crash test ratings received considerable news media and public attention (Table 5). In addition to traditional television news, IIHS's YouTube channel also attracted large numbers of viewers. Table 6 shows the number of views for the YouTube videos associated with each of the public releases of small overlap crash test ratings. The combined number of views for all small overlap crash test videos, including individual crash tests, exceeded 37 million as of mid-December 2014.

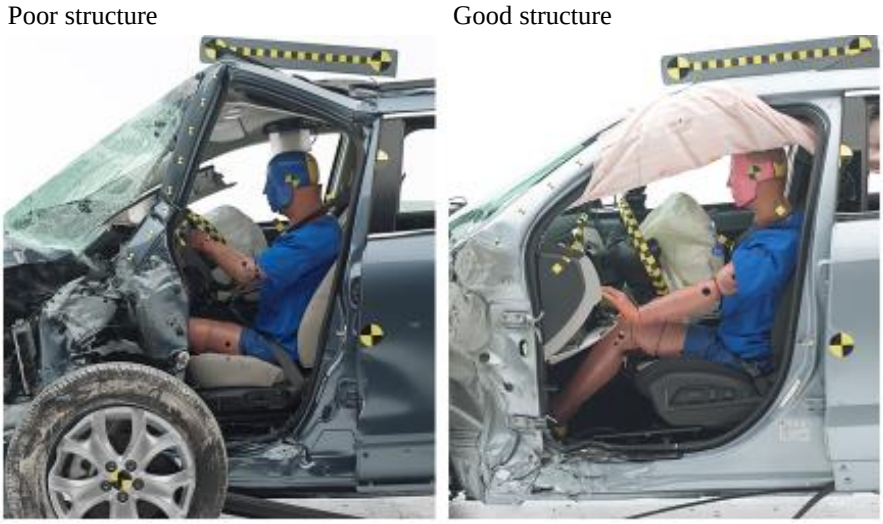


Figure 4. Comparison of survival space in Mazda CX-9 (poor) and Chevrolet Equinox (good) following a 64 km/h 25 percent overlap crash against a rigid barrier.

Figure 5. Comparison of driver dummy motion during a 64 km/h 25 percent overlap crash against a rigid barrier.

All of this attention to the new crash test ratings has also elevated consumer interest in vehicles earning good ratings, at least in the short term. Surveys of vehicle dealerships indicated that announcements of crash test ratings increased consumer interest in and sales of good-rated models in the week following ratings announcements compared with the week before [19]. Thus, it seems worth automakers' efforts to design their products to earn good crash test ratings.

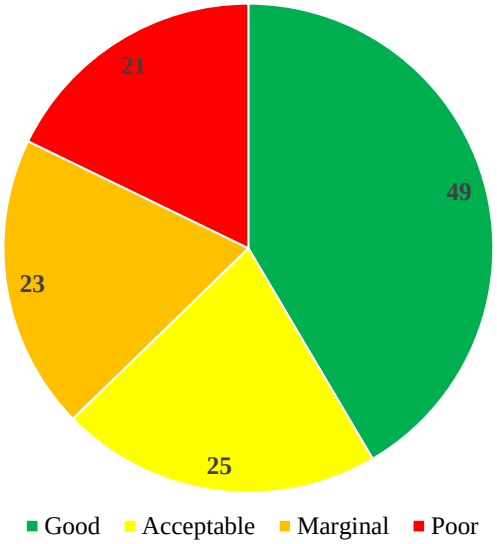


Figure 6. Small overlap front crash test ratings for 2015 model year vehicles.

Table 5.
Number of television broadcasts and estimated audience for IIHS front small overlap crash test ratings announcements.

Date	Vehicle group	Number of broadcasts	Estimated audience (millions)
November 20, 2014	Minivans	1,905	129.2
July 30, 2014	Small cars 2	3,245	165.5
April 8, 2014	Midsize SUVs	2,441	146.2
January 22, 2014	Minicars	2,048	121.9
August 8, 2013	Small cars 1	1,129	110
May 16, 2013	Small SUVs	790	63.4
December 20, 2012	Midsize non-luxury cars	1,262	155.5
August 14, 2012	Midsize luxury cars	2,751	212.3

Table 6.
Number of views for IIHS small overlap crash test ratings news videos on YouTube, as of January 23, 2015.

Date	Vehicle group	Number of YouTube views
November 20, 2014	Minivans	209,845
July 30, 2014	Small cars 2	174,944
April 8, 2014	Midsize SUVs	229,127
January 22, 2014	Minicars	426,338
August 8, 2013	Small cars 1	179,066
May 16, 2013	Small SUVs	235,290
December 20, 2012	Midsize non-luxury cars	951,370
August 14, 2012	Midsize luxury cars	816,218

Front Crash Prevention Systems Ratings

HLDI published a series of analyses using insurance data that evaluated the benefit of new crash avoidance technologies [20][21][22]. These analyses showed that all of the systems intended to help drivers avoid being in front-to-rear collisions were preventing some crashes reported to insurers. Vehicles with forward collision warning (FCW) from three automakers had 7-10 percent fewer liability claims for damage done to other vehicles and 4-7 percent fewer collision claims for own damage compared with the same year/make/models without FCW. The frequency of injury claims also was reduced. The results for automatic braking systems were even better. Volvo vehicles with City Safety (a standard feature) have 15 percent fewer liability claims and 18 percent fewer collision claims than their market competitors without automatic braking systems, and injury claims are reduced by more than 25 percent. Similarly, optional automatic braking systems from four different automakers also reduce liability claims compared with their counterparts without it. There is an indication that own damage and injury claims also are reduced with these systems, but the confidence intervals for the estimated reductions are large.

Based on these crash and injury reductions and consistency across different automakers' implementations of FCP systems, IIHS began rating them in 2013. The ratings — basic, advanced, and superior — are intended to reflect the level of benefit consumers could expect from systems with different characteristics. The basic level is intended to identify models available with FCW, although it is possible to earn the basic rating with automatic braking that produces small speed reductions in one of two simulated front-to-rear crashes. The FCW performance requirements are the same as those that the National Highway Traffic Safety Administration uses to identify FCW in its New Car Assessment Program. The advanced rating is intended to represent models with the availability of automatic braking systems similar to Volvo's City Safety or the optional systems studied by HLDI. Typically this requires collision prevention or near prevention through automatic braking in one of two simulated front-to-rear collisions at 20 and 40 km/h. However, it is possible to earn the advanced rating with moderate speed reductions in both tests. Models that combine FCW with automatic braking earn more points toward their ratings than those with automatic braking alone. The superior rating is intended to identify models with more capable automatic braking and requires that both simulated collisions be avoided or nearly so through the automatic application of brakes. Figure 7 shows that the availability of these systems at every rating level has increased in the 2015 model year compared with 2013.

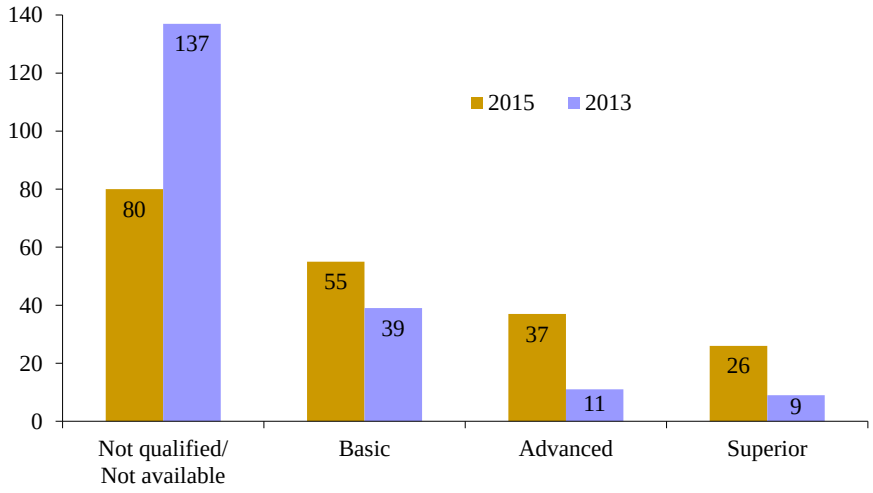


Figure 7. Comparison of the availability of FCP systems with different IIHS ratings, 2015 vs. 2013 model year

IIHS FCP ratings were first published in September 2013, and an update was published in May 2014. Both received a moderate amount of news and public attention, with estimated audiences of 58 and 79 million, respectively, in addition to being highlighted in *Top Safety Pick* announcements since 2013. Consumer response, as measured by reported level of interest and sales at automobile dealerships, was muted in comparison with what was observed for

crashworthiness ratings [19]. Nevertheless, compared with dealerships whose marquis brands did not offer FCP technology on any product, those that were mentioned in the FCP ratings announcements had better new vehicle sales experience.

FUTURE CONSUMER INFORMATION PROGRAMS

Vehicle LATCH Evaluations

All passenger vehicles sold in the United States have been required to be equipped with Lower Anchors and Tethers for Children (LATCH) since September 1, 2002 [23]. The purpose of LATCH, which is similar to ISOFIX, is to facilitate the proper installation of child restraint systems (CRS) and to provide top tether anchors for CRS installed with vehicle seat belts. Since its earliest implementation, however, many LATCH systems have failed to achieve this basic function, and observations of child restraint installations indicate that only slightly more than half of CRS in vehicles equipped with LATCH were installed using it [24][25]. Use of the top tether on forward-facing CRS is similarly low [26][27].

The Society of Automotive Engineers (SAE) Child Restraint Systems Subcommittee and ISO TC22/SC12/WG1/TF2 have drafted procedures and tools for assessing LATCH usability and the compatibility between vehicles and child restraints when using LATCH [28][29]. IIHS researchers working with UMTRI have established that the measures recommended by SAE and ISO do predict correct use of LATCH to install CRS [30][31][32][33]. Clear space around the anchor measured as an angle in a vertical plane, a low force to attach a simulated CRS LATCH connector, and a shallow depth of the anchor within the seat bight all were associated with higher rates of correct CRS installation using LATCH. This research will be the basis of a LATCH evaluation that IIHS intends to begin publishing in 2015. The evaluation will emphasize the ease-of-use characteristics in seating positions that are required by regulation to be equipped with LATCH and offer additional credit toward the highest rating for additional seating positions also equipped with easy-to-use LATCH.

In January 2015, NHTSA published a notice of proposed rulemaking that announced its intention to require easier-to-use LATCH [34]. The proposed requirements are consistent with IIHS and UMTRI's research, so the new LATCH evaluation will serve as a guide to parents purchasing new vehicles until all passenger vehicles are required to be equipped with better LATCH.

Advanced Head-lighting Evaluations

HLDI's studies of optional crash avoidance systems also found a significant benefit of adaptive headlights that was consistent across four automakers' implementations [20]. The frequency of property damage liability claims was reduced by 5-10 percent and collision claims were reduced by 1-6 percent for vehicles equipped with headlights that point in the direction the car is being steered, compared with their counterparts with traditional fixed-aim headlights. Large reductions of injury claim frequencies were also associated with steerable headlight systems.

The demonstrated benefits of these new headlight systems has inspired IIHS researchers to begin developing an evaluation of headlight systems to promote, through consumer recommendations, those that provide a better view of the road ahead. The main focus of this effort is identifying systems that enhance forward view on curved roads because this is the situation in which the steerable headlights studied by HLDI offer the greatest benefit over fixed-aim systems [35].

Ideally, the evaluation will not specify the technology by which enhanced lighting is achieved, so work has concentrated on making measurements of illumination as the tested vehicles are driven toward sensors placed on the test track (Figure 8). Curved paths are deemed necessary to assess how well each system illuminates curved roads. The curve radii will be based on locations of nighttime fatal crashes. Likely, tests will also include measurements

from straight approaches to the sensors because more injurious and fatal crashes at night occur on straight roads than on curves (Figure 9). Furthermore, initial measurements indicate a large range in straight-line distance that different headlight systems illuminate the road (Figure 10). The initial evaluations of new car headlight systems are expected to be published in 2015 with possible integration into the *Top Safety Pick* program in 2017.

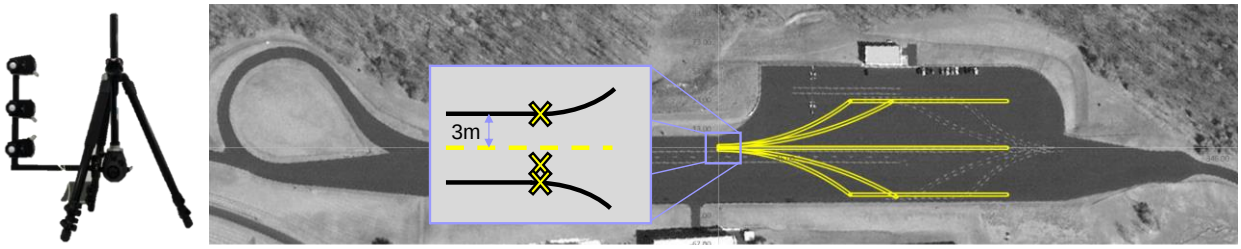


Figure 8. Conceptual set-up for IIHS headlight evaluations.

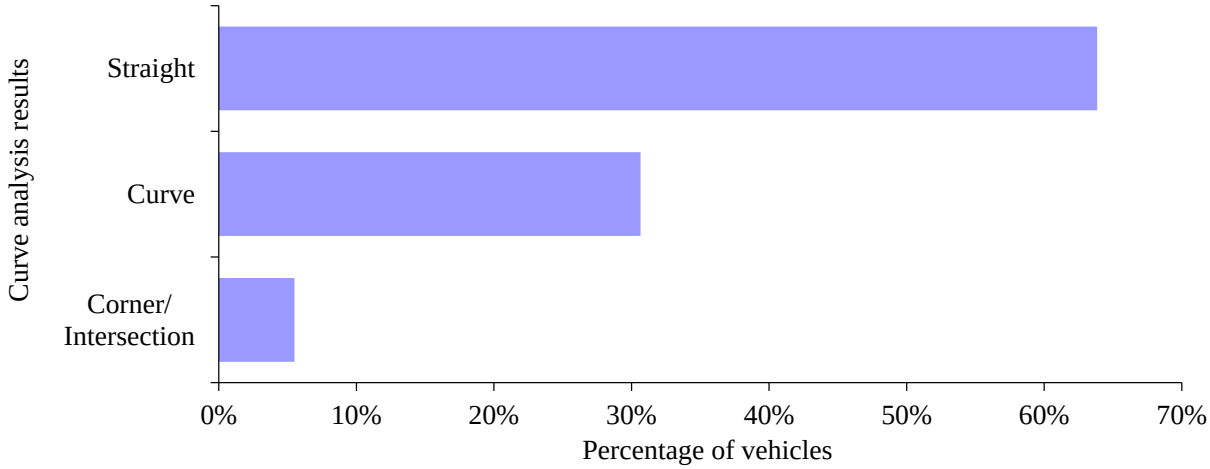


Figure 9. Distribution of road types for nighttime fatal crashes based on 2012 Fatal Analysis Reporting System.

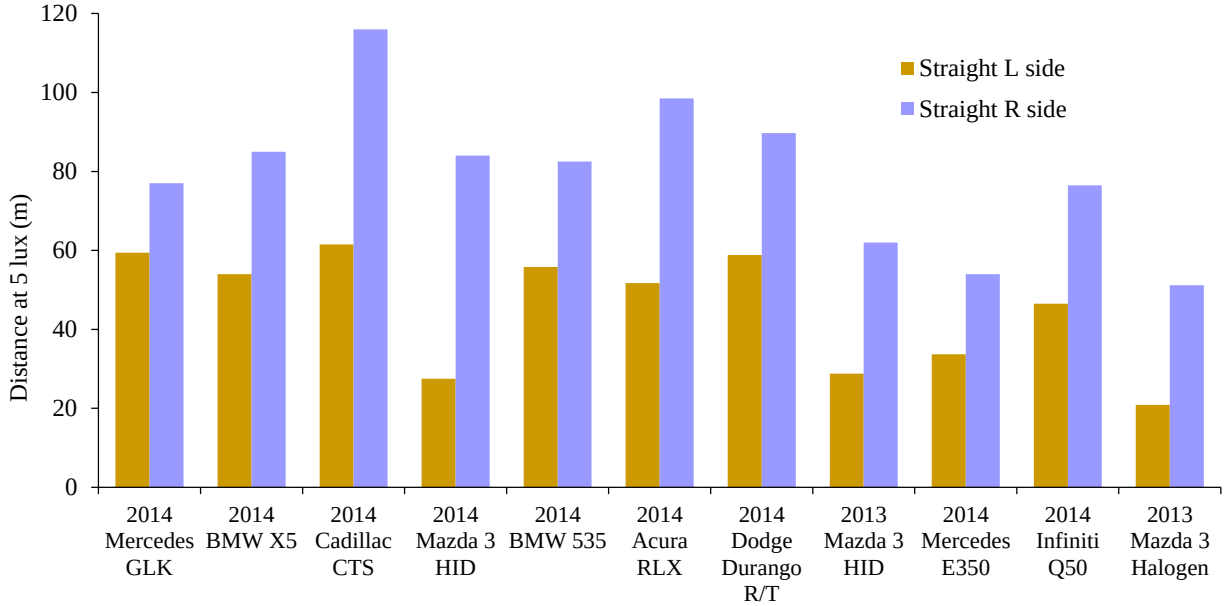


Figure 10. Maximum straight-line distance for 5 lux illumination at 25 cm above road level by year/make/model.

CONCLUSIONS

IIHS and HLDI have a long history of publishing comparisons of the crash safety of different vehicles. In some cases, this effort has led to legislative requirements for the distribution of information or regulatory requirements for the design of vehicles. More recently, large audiences for IIHS consumer information programs have prompted manufacturers of rated products to make changes in ways indicated by IIHS tests, and models achieving better ratings have been rewarded by increased sales. IIHS experience with truck underride guard testing also suggests that the consumer information model can promote improved safety in the commercial vehicle fleet. Based on its experience with current programs, there is good reason to believe that IIHS ratings of LATCH and advanced head lighting systems can also improve vehicle safety.

REFERENCES

[1] Hershman LL. 2001. The U.S. New Car Assessment Program (NCAP): past, present and future. Paper presented at the 17th International Technical Conference on the Enhanced Safety of Vehicles. Washington, DC: National Highway Traffic Safety Administration.

[2] Highway Loss Data Institute. 1973. Automobile insurance losses, collision coverages: variations by make and series, 1972 models. Washington, DC.

[3] Motor Vehicle Information and Savings Act, 92 P.L. 513, 86 Stat. 947 (1972)

[4] Repeal of Requiring Motor Vehicle Insurance Cost Reporting, 112 P.L. 252, 126 Stat. 2406 (2013)

[5] Farmer CM. 2005. Relationships of frontal offset crash test results to real-world driver fatality rates. *Traffic Injury Prevention* 6:31-37.

[6] Teoh ER.; Lund AK. 2011. IIHS side crash test ratings and occupant death risk in real-world crashes. *Traffic Injury Prevention* 12:500-07.

[7] Farmer CM, Zuby DS, Wells JK, Hellinga LA. 2008. Relationship of dynamic seat ratings to real-world neck injury rates. *Traffic Injury Prevention* 9:561-67.

[8] Brumbelow ML, Teoh ER, Zuby DS, McCartt AT. 2009. Roof strength and injury risk in rollover crashes. *Traffic Injury Prevention* 10:252-65.

[9] Brumbelow ML, Teoh ER. 2009. Roof strength and injury risk in rollover crashes of passenger cars. *Traffic Injury Prevention* 10:584-92.

[10] Arbogast KB, Jermakian JS, Kallan MJ, Durbin DR. 2009. Effectiveness of belt positioning booster seats: an updated assessment. *Pediatrics* 124:1281-86.

[11] O'Neil J, Daniels DM, Talty JL, Bull MJ. 2009. Seat belt misuse among children transported in belt-positioning booster seats. *Accident Analysis and Prevention* 41:425-29.

[12] Reed MP, Ebert SM, Sherwood CP, Klinich KD, Manary MA. 2009. Evaluation of the static belt fit provided by belt-positioning booster seats. *Accident Analysis and Prevention* 41(3):598-607.

[13] Braver ER, Cammisa MX, Lund AK, Early N, Mitter EL, Powell MR. 1997. Incidence of large truck-passenger vehicle underride crashes in the Fatal Accident Reporting System and the National Accident Sampling System. *Transportation Research Record* 1595:27-33.

[14] Brumbelow ML, Blanar L. 2010. Evaluation of U.S. rear underride guard regulation for large trucks using real-world crashes. *Stapp Car Crash Journal* 54:119-131.

[15] Insurance Institute for Highway Safety. 2011. Crash tests demonstrate the need for new underride guard standards. *Status Report* 46(2). Arlington, VA.

[16] Insurance Institute for Highway Safety. 2013. Not good enough: Underride guards on big rigs can be lifesavers, but most leave passenger vehicle occupants at risk in certain crashes. *Status Report* 48(2). Arlington, VA.

[17] Brumbelow ML, Zuby DS. 2009. Impact and injury patterns in frontal crashes of vehicles with good ratings for frontal crash protection. *Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles* (CD-ROM). Washington, DC: National Highway Traffic Safety Administration.

[18] Mueller BC, Brethwaite AS, Zuby DS, Nolan JM. 2014. Structural design strategies for improved small overlap crashworthiness performance. *Stapp Car Crash Journal* 58:145-74.

[19] Cicchino J. 2015. Consumer response to vehicle safety ratings. *Proceedings of the 24th International Technical Conference on the Enhanced Safety of Vehicles* (CD-ROM). Washington, DC: National Highway Traffic Safety Administration.

[20] Moore M, Zuby D. 2013. Collision avoidance features: initial results. *Proceedings of the 23rd International Technical Conference on the Enhanced Safety of Vehicles* (CD-ROM). Washington, DC: National Highway Traffic Safety Administration.

[21] Insurance Institute for Highway Safety. 2014. Honda warning system trims insurance claims. *Status Report* 49(4). Arlington, VA.

[22] Insurance Institute for Highway Safety. 2013. First crash avoidance ratings under new test program: 7 midsize vehicles earn top marks. *Status Report* 48(7). Arlington, VA.

[23] Office of the Federal Register. 2011. *Code of Federal Regulations*; Title 49 Transportation, Part 571 Federal Motor Vehicle Safety Standards, Standard No. 225 Child Restraint Anchorage Systems. Washington, DC: National Archives and Records Administration.

[24] Insurance Institute for Highway Safety. 2003. LATCH kids in cars. *Status Report* 38(5). Arlington, VA.

[25] Decina LE, Lococo KH. 2007. Observed LATCH use and misuse characteristics of child restraint systems in seven states. *Journal of Safety Research* 38:273-81.

[26] Eichelberger AH, Decina LE, Jermakian JS, McCartt AT. 2014. Use of top tethers with forward-facing child restraints: observations and driver interviews. *Journal of Safety Research* 48:71-76.

[27] Jermakian JS, Wells JK. 2011. Observed use of tethers in forward-facing child restraint systems. *Injury Prevention* 17:371-74.

[28] Society of Automotive Engineers. 2009. Guidelines for implementation of the child restraint anchorage system or LATCH system in motor vehicles and child restraint systems. Document no. J2893. Child Restraints Systems Committee, Troy, MI. Warrendale, PA.

[29] International Organization for Standardization. 2010. Road vehicles: Methods and criteria for usability evaluation of child restraint systems and their interface with vehicle anchorage systems; Part 1 Vehicles and child restraints equipped with ISOFIX anchorages and attachments. Report no. CD 29061-1. Geneva, Switzerland.

[30] Klinich KD, Manary MA, Flannagan CAC, Moore JL, Jermakian JS. 2013. Evaluation of proposed protocols for assessing vehicle LATCH system usability. *SAE International Journal of Transportation Safety* 1:32-45.

[31] Klinich KD, Flannagan CAC, Jermakian JS, McCartt AT, Manary MA, Moore JL, Wells JK. 2013. Vehicle LATCH system features associated with correct child restraint installations. *Traffic Injury Prevention* 14:520-31.

[32] Jermakian JS, Klinich KD, Orton NR, Flannagan CAC, Manary MA, Malik LA, Narayanaswamy P. 2014. Factors affecting tether use and correct use in child restraint installations. *Journal of Safety Research* 51:99-108.

[33] Cicchino JB, Jermakian JS. 2014. Vehicle characteristics associated with LATCH use and correct use in real-world child restraint installations. Arlington, VA: Insurance Institute for Highway Safety.

[34] Office of the Federal Register. 2015. *Federal Register*, vol. 80, no. 15, pp. 3744-3785. National Highway Traffic Safety Administration, Notice of Proposed Rulemaking, Docket no. NHTSA-2014-0123; 49 CFR Part 571, Federal Motor Vehicle Safety Standards, Child Restraint Systems, Child Restraint Anchorage Systems. Washington, DC: National Archives and Records Administration.

[35] Reagan IJ, Brumbelow ML, Frischmann T. 2015. On-road experiment to assess drivers' detection of roadside targets as a function of headlight system, target placement, and target reflectance. *Accident Analysis and Prevention* 76:74-82.

The Australian Child Restraint Evaluation Program: An Update on Dynamic Assessment Protocols and Observed Dynamic Performance of Child Restraint Systems

Julie Brown
NeuRA Injury Prevention Research Centre and University of NSW
Australia

Basuki Suratno
Dan Leavy
Transport for NSW
Australia

Paul Kelly
Michael Griffiths
Road Safety Solutions
Australia

Peter Aquilina
Roads and Maritime Services, Crashlab
Australia

Michael Paine
Vehicle Design & Research
Australia

Jack Haley
NRMA Motoring and Services
Australia

Blake Harris
RACV
Australia

Paper Number 15-0376

ABSTRACT

The purpose of the Australian Child Restraint Evaluation Program (CREP) is to provide consumers with independent safety information; and to apply commercial and public consumer pressure on manufacturers to deliver child restraint systems (CRS) that perform well beyond the requirements of the Australian Standard. This paper describes the evolution of the dynamic assessment protocols and presents a summary of areas where improvement in dynamic performance has occurred. Areas of dynamic performance where there is still room for improvement, are also reviewed.

The dynamic assessment protocol has evolved from a system that separately scored the performance of CRS in frontal, 90 degrees and 66 degrees simulated impacts to a system that provides a single overall score for front and side impact tests to determine CRS ratings. The current protocols also nominate a number of 'Critical' 'Performance Aspect' (PAs) and a CRS is limited to one star if a score of '0' is achieved for any critical PA. There have also been significant changes to the dynamic test and assessment methods over the years to ensure assessment methods are as objective as possible, and some variation in the types of performance features assessed. For rearward facing infant restraints, CREP currently assesses the ability of the CRS to retain the head and torso in front and side impacts, control upward and rotational displacement of the CRS in rebound and distribute the load over the back of the dummy, in frontal testing, , manage dummy head and torso energy in frontal testing and manage dummy head energy in side impact. Similar assessments of dummy and head retention and energy management are used in the rating of forward facing child restraints. These assessments also include head and knee excursion. For booster seats, the ability of the booster to provide and maintain good sash belt geometry, and to prevent submarining in frontal

impacts. Assessments of head retention and energy management in side impact and dummy retention both in near and off-side impacts are also included for booster seats. There have been substantial improvements in the side impact protection features of rearward facing and forward facing child restraints observed in the program, and increasingly better performance of booster seats in maintaining good seat belt geometry in frontal impact. However, there is a need for more attention to head energy management in side impact, particularly among rearward facing restraints. Among rearward facing restraints, there are also concerns about poor performance of most restraints to adequately distribute crash forces through the back of the torso in frontal impact. Among forward facing restraints, there are concerns over head containment during rebound in frontal impact.

While there have been significant improvements to the test and assessment methods used in CREP there is a possibility that some aspects of good performance are being overstated and aspects of poor performance understated due to limitations in the assessment and rating procedures. Areas for possible future refinements of the protocols are also discussed.

INTRODUCTION

A child restraint evaluation program (CREP) has been operating in Australia since 1994. The concept of the program has been described previously (Kelly et al, 1996). According to Kelly et al (1996) the program was introduced following laboratory crash test observations of substantial variations in the performance of child restraint systems coming onto the Australian market in 1993. All child restraint systems sold in Australia since 1978 have had to comply with Australia/New Zealand Standard 1754 (AS/NZS 1754 had). The Australian Standard sets a minimum level of required performance and the relatively small number of restraints on the Australian market prior to 1993 generally exceeded this level of performance in the laboratory and in the field. The observation that some new restraints from new manufacturers that were coming on the market appeared to be just meeting the minimum requirements of the Standard raised concerns that these restraints may not provide the same level of protection to children in the real world (Kelly et al, 1996). The very first iteration of the program (CREP-Stage 1) included additional tests and performance criteria to those prescribed by the Standard to establish any differences in performance beyond the minimum requirements of the Standard of restraints being sold in Australia at that time , and to bring these differences to the attention of consumers.

The results of Stage 1 were published in a national subscription consumer magazine and wider consumer access was facilitated through a publically available brochure. According to Kelly et al (1996) there was clear consumer interest in the program. It also became clear that the published information was being used by consumers as a tool in making purchasing decisions with manufacturers reporting increases in sales of restraints that ranked well, and decreases in sales of those that were rated poorly. This established the value of the program as a mechanism to apply commercial and public consumer pressure on manufacturers to deliver child restraint systems (CRS) that performed well beyond the requirements of the Australian Standard.

Two further releases of CREP results occurred in 1996 (Stage 2) and 2000 (Stage 3). These used the same general test methods, approach to assessing performance and publication strategies used in Stage 1.

In 2004 the program underwent a major review that resulted in a revised dynamic test protocol and a new approach to the overall assessment of the restraints using an objective point based method (Brown et al, 2007). Results of assessments using these new protocols were released as CREP Stage 4a & CREP Stage 4b. The assessment method was further enhanced in 2009, and the results from Stages 4A to 4D were rescored to the enhanced method and released in 2010, together with the results for CREP 4E.

Following the introduction of substantive changes to the Australian Standard in 2010 the test protocols were again reviewed in 2012 to ensure CREP stayed true to its original aim of assessing the performance of restraints sold in Australia beyond the minimum requirements of the Australian Standard. This resulted in more significant changes to the test method and inclusion of some additional performance assessments. The resulting test procedures and assessment protocols were adopted and have been used in the CREP 5 ratings. The CREP 5 rating procedures remain current with some slight modifications to allow for the assessment of ISOFIX compatible restraints.

This paper describes the evolution of the dynamic assessment protocols and presents a summary of the areas where improvement in dynamic performance has occurred. Areas of dynamic performance where there is still room for improvement are also reviewed.

EVOLUTION OF THE DYNAMIC ASSESMENT PROTOCOL

Throughout the first three releases of CREP results (CREP1-3) the assesment procedure remained relatively stable, however significnat changes were introduced in CREP 4 and CREP 5. These are summarised below.

Test Methods

While the intention of CREP since its inception has been to assess the performance of restraint systems beyond the minimum requirements of the Australian Standard, the test methods have been heavily based on the test methods included in the Standard. The test orientations and test pulses used in each iteration of the program are summarised in Table 1.

Table1.
Evolution of crash tests included in CREP assessment procedures

Sled Test	CREP Stage 1-2 (1994-1996)	CREP Stage 2 (1996-1999)	CREP Stage 3 (2000-2004)	CREP Stage 4 (2005-2012)	CREP Stage 5 (2012-present)
Frontal 48km/h 24g	RF, FF, Boosters	RF, FF, Boosters	RF, FF, HBB	-	-
Frontal 56km/h 34g	RF, FF, Boosters	RF, FF, Boosters	RF, FF	RF, FF, Boosters	RF, FF, Boosters
NS Side 90° +	RF, FF, Boosters	RF, FF, Boosters	RF, FF, HBB	RF, FF, Boosters	RF, FF, Boosters
NS Side 45° +	RF, FF, Boosters	RF, FF, Boosters	RF, FF, HBB	-	-
NS Side 66 ° +	-	-	-	RF, FF, Boosters	RF, FF
FS Side 90 ° +	-	-	-	-	Boosters
Rear Impact+	RF, FF	RF, FF	RF, FF	-	-
Inverted++	RF	RF	RF	-	-

*RF – Rearward facing infant restraint; FF – Forward facing child seat; LBB – Low back/backless booster; HBB – High back booster; Booster – any booster seat; NS – Near side; FS – Far side.
+ Impact pulse is 32km/h 14g and static side door structure incorporated in test set up; ++ Impact pulse 16km/h 8g

Initially (Stages 1-3) CREP included two frontal impact tests; one that equaled the minimum deceleration and velocity change requirements of the Australian Standard (48km/h, 24g), and one that attempted to more closely match the assessment of adult restraint systems in the Australian New Car Assessment Program (Kelly et al, 1996). The more severe pulse used a velocity change of 56km/h and a deceleration that was as near as possible to the maximum acceleration allowed by the Australian Standard test method (34g). However in the review of the protocols prior to CREP 4, the lower severity test was removed from the protocol as this test was not producing any worthwhile information for making comparisons of performance beyond that required by the Standard.

The side impact tests included in all iterations of CREP have used the same pulse requirements as that included in the Australian Standard. The Australian Standard has required a 90 degree side impact test since 1975. However, the assessment of performance in side impact in the Standard was originally related to the restraint's ability to retain the dummy and maintain structural integrity in this orientation. To allow restraint performance to be assessed beyond this, a side door structure was added to the test set up in CREP Stage 1 and as described below, this continues to CREP 5 with some modification to the structure. The additional 45 degree side impact test was included in CREP Stage 1 in acknowledgement of field data that indicated that most side impacts included some forward component of force. In the review conducted prior to CREP 4 it was felt that there was also no worthwhile comparable information being collected from the 45° tests, and it was unclear what proportion of real world impacts this orientation was actually simulating. To maintain the intention of being

able to assess performance under oblique loading but with some justification for the orientation being used, a decision was made to move to a 66° impact as that more closely resembled the US NCAP side impact conditions at that time.

Up until CREP Stage 5, all side impact testing simulated near side impacts. In the early stages of CREP 1-4, many high back boosters demonstrated inadequacies in the retention of test dummies during rebound in these near side impacts. To address this and allow the performance of different booster seats on the market to be compared, a far side impact was included for booster seats in CREP 5.

Rear and inverted tests included in CREP Stages 1-3 were basically mimicking the tests carried out as part of compliance testing for the Standard and were not providing any information on performance beyond that required by the Standard. These tests were discontinued after CREP Stage 3.

As shown in Table 1, there has been some variation in the restraint types subjected to different tests throughout the evolution of the program. As described by Brown et al (2007), all restraints were subjected to all frontal and side impact tests in the first Stage of CREP (Stage 1). The performance of most boosters on the market in Australia at that time (early 90's) was quite poor in that first series of tests. Based on this poor performance, subsequent early CREP series treated boosters differently. High back boosters were exempted from the higher severity frontal tests due to this poor performance and test house concerns regarding the robustness of the test TNO P10 dummy in CREP Stages 2 & 3. However trials conducted prior to CREP Stage 4 alleviated these concerns and these restraints were again subjected to the higher severity frontal impact in later stages. Booster cushions (i.e. backless boosters) were excluded from CREP Stages 2 and 3 due to their inherent inability to provide any measurable protection in side impact. This was reversed in the lead-up to CREP Stage 4, as changes made to the Australian Standard prior to the commencement of CREP 4 basically removed booster cushions or backless boosters from the Australian market. As described below, the change to the Australian Standard that impacted booster cushions/backless boosters was the adoption by Standards Australia of the CREP side impact test method and assessment procedure. Therefore no backless boosters have been included in any CREP series since Stage 1.

Booster seats were never included in the rear impact or inverted tests, and forward facing restraints were always excluded from the inverted tests.

Test rig

The test equipment used throughout all stages of CREP is the same as that used in Standards testing of child restraints in Australia. This involves a test bench in compliance with the Australian standard mounted on a rebound sled. For side impact tests, a side door structure is also mounted to the sled. There has been some evolution of the side door structure over the course of the program and this is described below.

As described by Kelly et al (1996) in the original CREP (Stage 1) the side door structure attempted to replicate a simplified rear door of a large sedan in size and shape, and included an inner door skin and 'window'. The structure was fabricated from square wall tubing, the inner skin from thin sheet metal and the 'window' from 6mm Polycarbonate sheeting. The inner skin was replaced after each test. A side door structure, comprising a metal frame and polycarbonate panels, was added to the Australian Standard test procedure in 2004. The addition of this structure, together with a requirement that all restraints prevent contact between the test dummy's head and the side door meant that by CREP Stage 4 many restraints on the Australian market were able to contain the head to some extent. To this point, head accelerations and Head Injury Criteria (HIC) were being recorded in the CREP tests but not being used in scoring. Even allowing for the unknown biofidelic nature of the dummy's head, the generally very high HICs being recorded indicated a distinct lack of energy attenuating materials in CRS side structures. This led to discussion about the possibility of including some assessment of the head energy management in side impact. However, the non-uniform nature of the side door structure prevented this from being achieved. Research programs that had attempted to alter the energy absorption properties of restraint side wings were not able to measure significant differences in dummy head acceleration and HIC (Bilston et al, 2005), because altering the side wings resulted in differences in head or restraint impact location on the door. As the stiffness of the original door varied with location, any variations due to modification of the side wings were masked by changes in where the door was struck. To counter this, a new door structure with uniform stiffness was introduced in CREP Stage 5.

Dummies

The dummies used in CREP also conform to the requirements of the Australian Standard and largely belong to the TNO P series family of dummies. However, there have been changes to the types of dummies used to test different restraint types, and to how some dummies are have been used over the evolution of the program. Table 2 summarises the test dummies used by restraint type and impact condition across the history of CREP.

Table2.
Test dummies used in CREP assessment procedures

Restraint type	Impact	CREP Stage 1 (1994-1996)	CREP Stage 2 (1996-1999)	CREP Stage 3 (2000-2004)	CREP Stage 4 (2005-2012)	CREP Stage 5 (2012-present)
Rearward Facing Infant	Front	P3/4 (9kg)	P3/4 (9kg)	P3/4 (9kg)	P3/4 (9kg) or P1 ½ (11kg)*	P3/4 (9kg) or P1 ½ (11kg)*
	Side	P3/4 (9kg)	P3/4 (9kg)	P3/4 (9kg)	P3/4 (9kg)	P3/4 (9kg)
	Rear/Inverted	Taru T# (4kg)	Taru T# (4kg)	Taru T# (4kg)	-	-
Forward Facing Child	Front	P6 (22kg)	P6 (22kg)	P6 (22kg)	P6 (22kg)	P6 (22kg)
	Side	P3 (15kg)	P3 (15kg)	P3 (15kg)	P3 (15kg)+	P3 (15kg)+
	Rear	P6 (22kg)	P6 (22kg)	P6 (22kg)	-	-
Booster (Type E)^	Front	P6 (22kg) or P10 (32kg)*	P6 (22kg) or P10 (32kg)*	P6 (22kg) or P10 (32kg)*	P10 (32kg)	P6 (26kg)++
	Side	P3 (15kg)	P3 (15kg)	P3 (15kg)	P6 (22kg)+	P6 (22kg)+
Booster (Type F)^	Front	-	-	-	-	P10 (32kg)
	Side	-	-	-	-	P10 (32kg)+

*Choice of test dummy depended on maximum occupant mass specified by manufacturer. +Seated height modified to better represent seated height of children at upper end of restraint size range (P3 boosted to 605mm, P6 increased by 40mm).++ P6 dummy weight is boosted to match upper mass range of Type E Booster. ^ A new type of booster was specified in the Australian Standard in 2010. # Taru Theresa.

The most important changes to the dummies occurred with the introduction of CREP 4 and were related to a realization that restraints on the market at that time (pre-2005) were often too small to accommodate the full size range of children to which they were marketed. This manifested, in many cases, as poor side impact protection, as the tops of the side wings of many restraints were below the seated height of the children using them. As shown in Table 2, the P3 dummy continued to be used to assess forward facing restraints, and the P6 to assess booster seats in CREP 4. However, the seated heights of these dummies were boosted to match the seated height of the upper age ranges of children using these types of restraints.

For booster seats, the size of the dummy used to assess performance in frontal impacts was reduced from the P10 in CREP 4 to the P6 in CREP 5. This occurred in response to manufacturers concerns that the P10 dummy is significantly bigger and heavier than the largest size of child for which these restraints were designed. There was evidence of this in CREP 4, with a few boosters observed to be unable to properly accommodate the larger dummy. Furthermore, since the primary objective of the booster is to improve seat belt geometry for small occupants who would not be adequately restrained by the belt without the booster, it was felt that the smaller P6 dummy would provide a more suitable assessment. However, the P6 dummy's weight is below that of the upper limit of the weight range of these restraints. Therefore, the P6 is weighted to match the upper limit of the mass range of Type E boosters, by adding material to the chest and pelvic areas. This is done to ensure there is good assessment of the structural integrity of these restraints in the frontal impact.

Performance assessment

The areas of performance assessed throughout the CREP history have remained relatively unchanged. Since its inception, the program has assessed a number of areas of performance common to all restraint types. These include structural integrity of the restraint and dummy retention. For rearward facing restraints there has also always been an assessment of the restraint's ability to distribute crash loads across the back of the dummy. 'The need to distribute the crash forces in this way arises from concerns about the potential for neck injury resulting

from axial loading of the spine, particularly in young infants with their combination of a relatively large head and weak neck. Head displacement and head energy management (HIC 36) in frontal impact has always been assessed in rearward facing and forward facing restraints and head protection in side impact has been common across all restraint types throughout the entire history of the program. However, there have been changes in what restraints are subjected to these assessments and the way head protection in side impact is monitored. There have also been other areas of performance assessment added for some restraints, over time.

Head displacement and head energy management were originally assessed in all restraint types but these assessments were eventually dropped for booster seats as these measures were found to be heavily influenced by lap belt and sash strap position. Low head excursion occurs when there is submarining. Although low head excursion is desirable, submarining in booster seats is not.

In the original Stage 1 of CREP, head protection in side impact was monitored in two ways. Firstly there was an assessment of whether or not the dummy's head was allowed to strike the side door structure, and secondly HIC 36 was used to gauge the severity of the impact between the dummy's head and the door structure. The head strike assessment has been continued through the subsequent stages of CREP to the present time, and retention requirement that restraints must prevent a head strike with the side door structure was included in the Australian Standard in 2004. In 2010, the Standard was revised to require that no observable part of the dummy head could be within 10mm of the door. No attempt was made to assess head injury management in side impact during the early stages of the CREP due to the non-uniform stiffness of the original door structure. However, HIC 36 values continued to be routinely collected to monitor the severity of head impacts and to gain an indication as to whether any meaningful energy attenuation measures were being implemented by manufacturers.

Following the introduction into the Australian Standard of the side door structure and the requirement for all restraints to prevent head to door contact, all devices coming onto the Australian market began to provide some degree of head containment. To capture the variability in the degree of head containment that was being observed, CREP Stage 4 introduced a graded score of this performance. As described by Brown et al (2007), the highest score is given if the head remains completely within the confines of the side wings and a low score is given if the head becomes exposed over the rim of the side wings. A '0' score is assigned to head contact with the door.

Assessment of side impact head energy management was introduced for CREP 5, using HIC 36 as the assessment criteria. This required the introduction of a new side door structure with uniform stiffness.

Significant additions to the assessment of booster seat performance were also made in CREP 4. As described by Brown et al, (2007) the assessment procedures introduced in CREP 4 placed a high priority on the ability of a booster seat to provide the test dummy with good seat belt geometry and for the dummy's upper torso to remain satisfactorily restrained throughout the impact. Assessments were therefore included for the pre-impact seat belt sash strap geometry on the dummy and for the maintenance or otherwise of good sash strap and lap belt geometries during impact. Similar assessments were subsequently incorporated into AS/NZS 1754:2010.

Scoring Systems

In CREP 1 and 2, "Preferred Buy" ratings were given to a number of devices in each category that were rated as performing well in a number of pre-defined areas. However, the method used to make these judgments was subjective and there was little documentation of the protocols used. This led to a significant difference in how preferred buys were awarded in CREP 3. In CREP 3 evaluations were based mainly on technical compliance with requirements, and little weight was given to the more desirable features of performance used in making assessments in CREP 1 & 2.

To address this, an objective point scoring system for CREP was first introduced in CREP 4. The scoring system used in Stage 4 involved the use of a ratings matrix, using an approach derived from methods used in the ease of use assessment by the North American National Highway Safety Traffic Administration (NHTSA), and is discussed below.

Basically, the features assessed in the dynamic component were divided into a set of performance categories. Within each category, there were a set of items or individual performance aspects (PA). Each PA was given a weight between 1 and 4 based on the importance in terms of offering crash protection in the real world. A numerical scale of 4 (good) to 0 (unacceptable) was used to rate the outcome for each PA. Scores for each PA

were obtained by multiplying the outcome score by the weight for that PA. Category scores were calculated by adding the scores obtained for that category and calculating what percentage this was of the maximum possible score for that category. Initially, each category was then awarded an A, B, C or D ranking based on the breakpoints set out in Table 3 below.

These breakpoints were set on the basis that any device scoring less than or equal to 50% of the maximum score was judged as 'unacceptable' and given a 'D' ranking. The range between 50% and 100% was then divided into 3 equal ranks.

The overall score was calculated by averaging the sum of the normalised scores for each component. That is [(% score Frontal) + (% score Side 66) + (% score Side 90)/3]. The overall score was then also awarded a ranking using the same breakpoints illustrated in Table 3. One limiting rule was also applied to category and overall rankings. This rule was that if any device receives two or more '0' scores (i.e. an 'unacceptable') score, that device could not be awarded an A or B ranking for that category or for an overall ranking.

Table 3 Ranking Score Calculations – CREP Dynamic Testing

'Performance Aspect' Set Score	Overall Score
The 'PA' set score ≥ 83% of maximum 'PA' scores that could be obtained for the set	The sum of the 'PA' scores ≥ 83% of the sum of the maximum scores that could be obtained for all the 'Performance Aspects'
The 'PA' set score < 83% but ≥ 67% maximum 'PA' scores that could be obtained for the set	The sum of the 'PA' scores ≥ 66% but < 83% of the sum of the maximum scores that could be obtained for all the 'Performance Aspects'
The 'PA' set score < 66% but ≥ 50% of maximum 'PA' scores that could be obtained for the set	The sum of the 'PA' scores ≥ 50% but < 66% of the sum of the maximum scores that could be obtained for all the 'Performance Aspects'
The 'PA' set score < 50% of maximum 'PA' scores that could be obtained for the set	The sum of the 'PA' scores < 50% of the sum of the maximum scores that could be obtained for all the 'Performance Aspects'

While the scoring system introduced in CREP 4 was relatively successful in providing comparative ratings of restraints on the market, some child restraint systems were observed to perform unexpectedly poorly in some important areas. Some of the observed failures were not anticipated when the CREP 4 protocols were introduced and the resulting ratings did not adequately reflect their poor performance. The lack of a 'balancing' feature in calculating the overall score from the component dynamic tests also added to the problem. Poor performance in one of the dynamic tests e.g. side impact could be masked by very good performance in the other tests e.g. frontal impact.

The scoring system was then modified and the revamped system introduced part way through CREP 4. It has not changed since. This new method has three important characteristics; 1. The assessment of performance in three different test orientations are combined into a single group of assessments; 2. A number of critical performance areas were identified where poor performance in one of those areas limits the overall score, and 3. A 5 star rating system replaces the 4 category A-D ranking system.

The method of calculating an overall score by averaging the sum of the normalized scores from each of the three dynamic tests was therefore abandoned and this addressed the need for any 'balancing' feature. As shown in Table 4, the three dynamic tests continued to be included and scored individually. However performance aspects that are common to two or more tests are grouped into a single score, with the lowest score in any test used in the ratings. Performance aspects that are not common across two or more tests are kept as separate items, with two exceptions. These are 'Head Retention' for forward facing seats and boosters; and dummy

retention in booster seats. While the assessment of 'Head Retention' for forward facing seats and boosters is common to both side impact tests, restraints often demonstrate very different performance in the two tests.

Table4.
Assessment features in the combined scoring system introduced during CREP4

Rearward facing infant restraints	Forward facing child restraints	Booster Seats
Head Retention - Front & Side Impact Testing*	Forward Head Excursion – Frontal Impact Testing*	Seat Belt Sash Strap Location - Frontal Testing
Dummy Retention - Front & Side Impact Testing*	Forward Knee Excursion – Frontal Impact Testing	Head Retention – 90 Side Impact Testing
Upward and/or Rotational Displacement of the CRS in Rebound – Frontal Impact Testing	Head Retention – Frontal and Side Impact Testing	Dummy Retention – Frontal Testing*
CRS Security and Integrity - Front & Side Impact Testing*	Dummy Retention - Front & Side Impact Testing*	Dummy Retention 90 Side Impact Testing
Load Distribution - Frontal Impact Testing	CRS Security & Integrity - Front & Side Impact Testing*	Head Energy Management -Side Impact Testing with the seat belt sash strap over the shoulder adjacent to the *Side Impact Door*
Head Energy Management - Frontal Impact Testing	Operation of Quick Release Device – Frontal Impact Testing*	Submarining - Frontal Impact Testing
Head Energy Management -Side Impact Testing*	Head Energy Management - Frontal Impact Testing	CRS Security & Integrity - Front & Side Impact Testing*
Torso Energy Management - Frontal Impact Testing	Head Energy Management - Side Impact Testing	
Adjuster Slip - Front & Side Impact Testing*	Torso Energy Management - Frontal Impact Testing	
	Upward Displacement of CRS in Rebound - Frontal Impact Testing	
	Adjuster Slip - Front & Side Impact Testing*	

Good performance in a 90 degrees test would almost certainly be negated by universally poor performance in the 66 degrees tests. Therefore, a decision was made to continue to treat these separately. Similarly, dummy retention in booster seats is quite often starkly different between frontal and side impact testing. Therefore, dummy retention in frontal impact continues to be treated differently from dummy retention in side impact.

In this new method of scoring the overall performance, the overall score is calculated by taking the sum of points achieved divided by the maximum number of possible points and expressing this as a percentage.

A number of performance aspects have been nominated as critical features to ensure extremely poor performance in any of the tests is adequately captured. If a restraint scores a 0 in any one of the critical features (highlighted with * in Table 4) the restraint is limited to a 1 star rating.

The five star rating system was achieved by dividing the range of scores between 50 and 100. This resulted in the following breakpoints:-

- 5 stars – ≥ 87.5%
- 4 stars – ≥ 75% < 87.5
- 3 stars - ≥ 62.5% < 75%
- 2 stars - ≥50% <62.5%
- 1 star - <50%

OBSERVED IMPROVEMENTS IN CHILD RESTRAINT SYSTEM PERFORMANCE

There have been a number of important and quite clear improvements in child restraint performance observed over the course of the CREP. The most important of these have been the improvements in side impact protection provided by forward facings seats and booster seats, and the performance of booster seats in frontal impact. As detailed below these have largely occurred following significant enhancements in restraint design.

Side impact protection

In the first series of the CREP (Stage 1), 40% (4/10) of forward facing restraints and 100% (5/5) of the high back booster seats, allowed a head strike. This was despite the fact that in this test series there was no increase in the seated height of the dummies to match the upper limit of the seated height range of children who would be using the restraints. An improvement in side impact protection is evident from review of the results coinciding with the beginning of CREP Stage 4 and the boosting of the dummy seated heights.

In the first series of CREP 4 (4a), most forward facing restraints and high back booster seats allowed head contact with the door. While this poor performance continued for most of CREP Stage 4 there was some improvement observed, see Figure 1. However, by CREP 5 no head contact with the side door occurred in any restraint in the 90° tests.

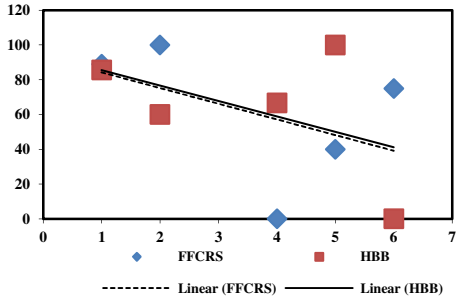


Figure1. Performance of restraints in each series of CREP 4 from CREP4a (1) to CREP 4B (2) shown as percentage of restraints that allowed head contact with side door in each series. FFCRS - Forward facing child restraint; HBB – High Back Booster

Typical side wing height change on FFCRS over CREP4-CREP5

Typical side wing height change on HBB over CREP4-CREP5

Figure2. Typical changes in side wing height in forward facing child restraints (FFCRS) and high back boosters (HBB) from CREP 4 (2005-2012) to CREP 5 (2012-)

As shown in Figure 2, the driver in these improvements is an increase in side wing height in both forward-facing child restraints, and high back booster seats. The movement towards higher side wings is likely to be a result of the combination of the adoption of the CREP side impact test procedure and a similar assessment method into the Australian Standard in 2004 together with additional requirements included in the Standard for minimum seat back heights for all restraint types in 2010. The restraint dimension requirements were written to support laws requiring appropriate restraint use by Australian children. Restraint back and (for rearward & forward facing restraints) shoulder harness slot heights were defined to ensure that forward facing restraints could accommodate children up to the 95th percentile 4 years old, and booster seats could accommodate children up to the 95th percentile 8 year old.

Frontal crash protection in booster seats

The ability of high back booster seats to provide good seat belt geometry and to maintain good torso restraint during frontal impact began to be assessed in CREP from Stage 4. Results from the first 17 high back booster seats from these assessments have been published previously (Brown et al, 2009). As described by Brown et al (2009), 10 of the 17 (59%) demonstrated static lap belt positions low down on the dummy abdomen, adjacent to the dummy thighs, and only 5 of the 17 (29%) boosters were able to maintain good lap belt positioning throughout the frontal impact. Furthermore, Brown et al (2009) noted a wide variation in the static position of the sash part of the belt in these boosters, with only two restraints positioning the sash over the mid-shoulder region of the dummy, and contact between the sash belt and the dummy shoulder was maintained during the impact in only 9 of the 17 boosters (53%).

The results presented by Brown et al (2009) represent the first half of the CREP 4 test series, and as shown in Figure 3, some improvement was beginning to be observed by the end of CREP 4. Fifteen boosters were tested in the second half of CREP 4, good lap belt positioning was maintained in 8 of the 15 boosters (53%) and contact between the sash belt and dummy shoulder maintained in 13/15 boosters (87%).

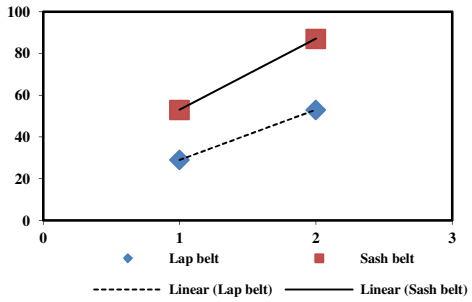


Figure3. Percentage of boosters maintaining good lap belt and sash belt position during dynamic frontal testing in the first half (1) and second half (2) of CREP 4 (2005-2012)

Requirements related to the static pre-impact position of the lap belt were included in the Australian Standard in 2010 – and more recently knee excursion limits have been added as a defacto measure of submarining. (2013). This means that all booster seats now sold on the Australian market must demonstrate the ability to provide good lap belt positioning and minimize submarining in frontal impact testing.

As described above, the dummy being used to assess most booster seats was changed to the P6 in CREP Stage 5. In CREP 5 there has only been one booster that has failed to maintain good lap belt position using the P6 dummy, and all have demonstrated an ability to maintain contact between the sash belt and the shoulder throughout testing.

AREAS IDENTIFIED FOR FURTHER IMPROVEMENT IN RESTRAINT PERFORMANCE

While there have been some notable improvements in performance as described above, some issues of concern remain. These include the level of side impact protection for the head and load distribution management in frontal testing in rearward facing restraints and head protection in forward facing restraints during rebound in frontal testing.

Head protection in side impact in rearward facing restraints

Head energy management assessment in side impact commenced in CREP 5. This involves the use of HIC36 to comparatively assess how well restraints attenuate energy once the dummy's head is contained within the side structure of the restraint in the 90° impact. As shown in Figure 4, there are some substantial variations being observed in how well different restraints appear to be able to do this. It is unknown whether the variations in performance being observed reflect any variation in performance in the real world, but the magnitude of the differences suggest this may be an area worthy of further investigation

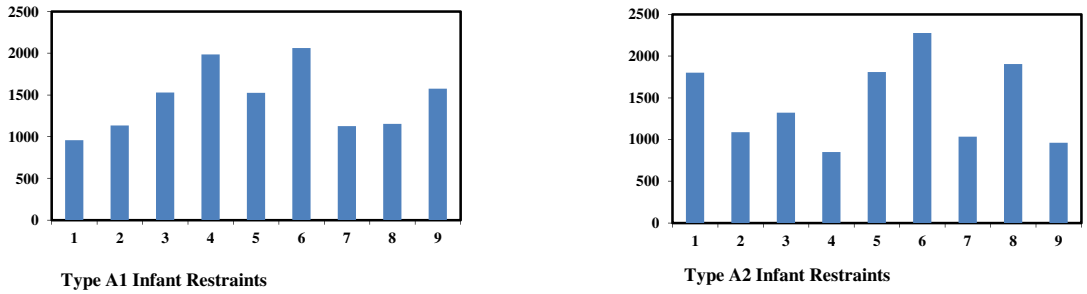


Figure4. HIC36 measured in 90° side impact with TNO P3/4 (Type A1) & TNO P1 ½ (Type A2)

Distribution of crash load in rearward facing restraints

Another area of concern in the performance of rearward facing restraints is their ability to distribute the crash load over the back of the dummy. A primary design goal of a rearward facing restraint should be to distribute most of the load in a frontal crash through the back of the head and torso. The Australian Standard carries a note to this effect, and as such it is not a mandatory requirement of the Standard. This feature has been assessed in CREP since its inception by calculating the force distributed through the back of the torso as a percentage of the force along the longitudinal axis of the torso. The larger the value, the better the restraint is in achieving the design goal. Figure 5 illustrates the results achieved by rearward facing restraints tested during CREP 5. There is clearly one restraint that is substantially superior in being able to distribute the greater proportion of the load through the back of the dummy. Significantly, 17 of the 27 (63%) of the restraints tested to dated in CREP 5 have not been able do so. Again the effect this is having in the real world is currently unknown, but the results indicate further investigation is warranted.

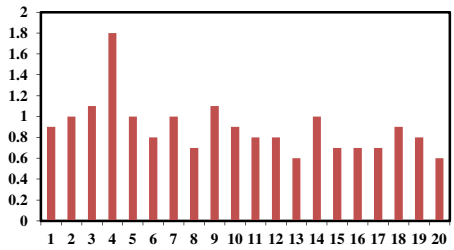


Figure5. Performance of rearward facing restraints in distributing crash loads over back of dummy head and torso in frontal impact (from CREP 5)

Head protection in forward facing restraints during rebound

While we have seen significant increases in the height of the side wings in forward facing restraints, and the full height of the side wings being comprised of structurally robust material, there appears to have been a tendency for manufacturers to not extend this height/structural integrity to the back of their restraints. In CREP 5, the behavior of the dummy's head during rebound in frontal testing has been assessed in forward facing restraints. This has captured a previously unreported potentially poor aspect of performance of many modern forward facing restraints. This involves the dummy's head being allowed to rotate over the top of the seat back (see Figure 6). This occurs in restraints where the height of the restraint back reduces between the side wings, as well as in restraints where the material at the top of the seat back fails to adequately support the dummy's head. The latter has commonly been seen in restraints incorporating internal adjustable head restraints (see Figure 6).

Head rotates overs 'dip' in seat back No structural support at top of back Head rest fails to support head

Figure6. Poor support of dummy head in rebound during frontal impact

AREAS IDENTIFIED FOR FURTHER ENHANCEMENT OF THE SCORING PROTOCOLS

On the introduction of an objective point system based assessment method in CREP 4, the focus for booster seats was primarily on the ability of boosters to achieve their primary design goal of improving seat belt geometry for the children using them. As described above, there have been substantial improvements in this regard in boosters introduced onto the Australian market over the last decade or so. There have also been significant enhancements to the Australian Standard related to the required levels of performance in maintaining good seat belt geometry in frontal impact. Changes made to the booster seat protocols prior to CREP 5 in response to the modifications made to the Australian Standard, may have left the protocols needing more refinement to better communicate variations in performance beyond that related to front impact protection.

Some boosters are performing in a far superior manner to others in containing the dummy's head in side impact and managing the energy, as measured by HIC36 (see Figure 7). Currently, the scoring system being used in CREP 5 is not capturing this variation in performance as well as it might. Five star boosters should be expected among the higher performing restraints in all aspects, and as shown in Figure 7, one restraint in particular has measured HIC >1000 in side impact but has still managed to score a 5 star rating. Similarly there are two restraints that have measured HICs among the highest values seen in this series and these restraints have still managed to score 4 stars. Further refinements of the assessment and scoring protocol are planned to try to adequately capture this type of performance in ratings. Moreover, consideration is being given to making the head energy management in side impact a critical feature. In that way this type of poor performance would limit the overall score the restraint could achieve. It is important that manufacturers be encouraged to seriously consider the energy attenuation provided by their restraint systems in side impact.

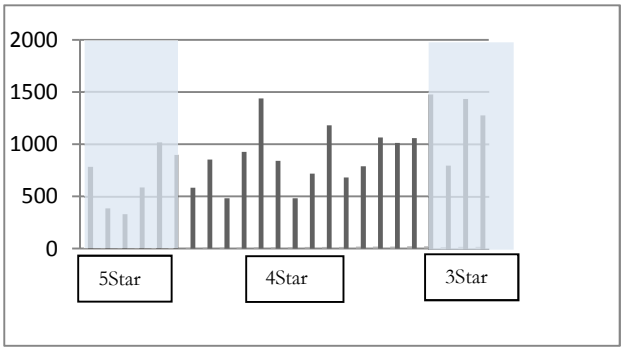


Figure7. Head energy management as measured by HIC3g in boosters in side impact in CREP 5

CONCLUSIONS

The Australian CREP program has evolved dramatically over the last 20 years. It is now a comprehensive and objective rating system providing important advice to consumers about the performance of restraints beyond the minimum performance level required by the Australian Standard for child restraints.

There have been innovative test methods and assessment procedures incorporated into the CREP since its inception and these have continued to evolve with the program. In addition to providing consumers with important information and encouraging manufacturers to continually improve their products, the CREP has also driven improvement in the minimum performance required by the Australian mandatory product Standard. Together, this has worked to influence some dramatic improvements in the performance of child restraints now available on the Australian market. This is most notably seen in the head protection provided in side impact in forward facing restraints and high back boosters; and the performance of high back boosters in frontal impact.

The program also continues to highlight emerging areas of concern. Currently this includes the head protection provided by rearward facing restraints in side impact, the inability of rearward facing restraints generally to adequately distribute crash forces in frontal impact, and the tendency for poor head containment in forward facing restraints during rebound in frontal testing. While it is clear that there are substantial variations in performance of restraints currently on the market in these areas, the impact of this on the real world crash protection of children remains unclear. Further investigation is warranted.

Finally, there may be some aspects of poor performance still not adequately being captured by the current protocols and work continues to further enhance the protocols.

REFERENCES

Bilston, Lynne E, Brown, Julie, & Kelly, Paul. (2005). Improved protection for children in forward-facing restraints during side impacts. *Traffic injury prevention, 6*(2), 135-146.

Brown, Julie, Kelly, Paul, Suratno, Basuki, Paine, Michael, & Griffiths, Michael. (2009). The need for enhanced protocols for assessing the dynamic performance of booster seats in frontal impacts. *Traffic injury prevention, 10*(1), 58-69.

Brown, Julie, Paine, Michael, Paine, David, Kelly, Paul, Griffiths, Michael, Magadera, Nimmi, Soames, J. (2007). *Revised assessment protocols and scoring methods for the Australian Child Restraint Evaluation Program.* Paper presented at the 20th Enhanced Safety of Vehicles Conference, Lyon.

Kelly, P, Griffiths, M, Booth, M, Lemon, J, Crothers, N, & Franks, C. (1996). *Child restraint evaluation program.* Paper presented at the Proceedings: International Technical Conference on the Enhanced Safety of Vehicles.