REVIEW OF EXISTING INJURY CRITERIA AND THEIR TOLERANCE LIMITS FOR WHIPLASH INJURIES WITH RESPECT TO TESTING EXPERIENCE AND RATING SYSTEMS

Klaus Bortenschlager
PDB – Partnership for Dummy Technology and Biomechanics, Germany
Markus Hartlieb
DaimlerChrysler AG, Germany
Karl Barnsteiner
BMW AG, Germany
Leonhard Ferdinand
Dr. Ing. h.c. F. Porsche AG, Germany
David Kramberger
Audi AG, Germany
Sven Siems
Volkswagen AG, Germany
Markus Muser
AGU Zürich, Switzerland
Kai-Uwe Schmitt
AGU Zürich, Switzerland, and
University and ETH Zürich, Switzerland
Paper Number 07-0486

ABSTRACT

In the recent years, a large effort has been directed towards the investigation of injury mechanisms and injury tolerance criteria related to whiplash associated disorders (WAD). Nevertheless, many questions, especially related to injury criteria and their respective biomechanical tolerance levels, remain unresolved. With the introduction of consumer tests in which the protection potential of seats against WAD is evaluated, a discussion of the criteria used for these ratings is needed, since for most proposed WAD injury criteria, e.g. NIC, Nkm, no widely accepted tolerance levels or even accurate injury risk curves are available today. One of the often disregarded points in the tolerance limit discussions is the fact that most injury criteria values have a non-linear relation to injury risk. Many tolerance levels for criteria related to injuries other than WAD (such as HIC, Nij, TTI, TI etc.) were derived using highly non-linear logistic regression curves. The biomechanical loads discussed in conjunction with WAD, e.g. accelerations, forces, torque, are generally very low in comparison to loads acting in other crash situations. Therefore, even minor changes in a test set-up may result in significant changes in the loads measured. Furthermore, issues of repeatability and reproducibility become more important in these low-load test conditions.
A series of sled tests was conducted to assess the influence of several test parameters on the repeatability of results obtained with the BioRID-IIg Dummy. The sled tests were performed according to the test procedure proposed by EuroNCAP. The results show that some criteria like the neck shear force exhibit variations up to 30%. The influence of such deviations has to be considered when introducing a reliable rating system for WAD.

INTRODUCTION

Soft tissue neck injuries as sustained in rear-end collisions are still a major concern in road traffic safety. Despite the various research that was undertaken in the last years the underlying injury mechanism is biomechanically not yet fully understood. Nevertheless, several injury predictors are proposed. Some of these show good correlations with real world accident studies and therefore seem suited to assess the injury risk. However, due to the complex nature of the injury, even for those criteria uncertainties remain with respect to the threshold values suggested. Since accurate injury risk curves are often not available, it is difficult to clearly define a threshold value which can reliably be regarded as a limit for injury.
Despite these uncertainties, there are indications that an improved seat design reduces WAD. Therefore, attempts are made to encourage manufactures to improve seat design. One way to achieve this, is the adoption of seat tests in consumer rating-programs. Consequently, evaluating the performance of seats with regard to WAD is currently widely discussed. Several studies showed that sled tests seem a suitable method to investigate the behaviour of a seat in low speed rear-impact [e.g. 1, 2]. Additionally, static measurements (geometric head-restraint assessment) as defined by several organizations aim in improving head restraint geometry.
In this study the repeatability of seat assessment using static and dynamic tests is investigated. Furthermore, by applying a rating scheme, the influence of data variations on rating results is demonstrated.

MATERIALS & METHODS

To evaluate different injury criteria, a series of tests consisting of static as well as dynamic tests (i.e. sled tests) were performed. One current car seat model from a high volume car manufacturer was chosen. All tests were conducted by using this seat model. This seat does not feature any (re-)active system to prevent WAD. For each sled test, a new seat was used, i.e. no seat was loaded twice. The head restraint was positioned identically for all tests (head restraint was locked in mid-height position) and the seat back angle was always adjusted to a 25° ±0.2° torso line measured by SAE-J826 H-Point Manikin.
A BioRID-IIg dummy of the latest build level was used throughout this study and certified prior and after the test series. The dummy was seated according to IIWPG procedure [3].

Static tests

Prior to each sled test, static measurements were taken to determine the head restraint height and the backset (i.e. the horizontal head to head restraint distance). The data was acquired and recorded as described in the IIWPG geometry measurement technique [3] using a SAE H-Point machine according to SAE-J826 equipped with the Head Restraint Measuring Device (HRMD).

Sled tests

Dynamic testing was performed using a HyperG220 acceleration-sled on which the seats were rigidly mounted. All seats were adjusted in the same way. The BioRID-IIg was instrumented according to IIWPG [3].

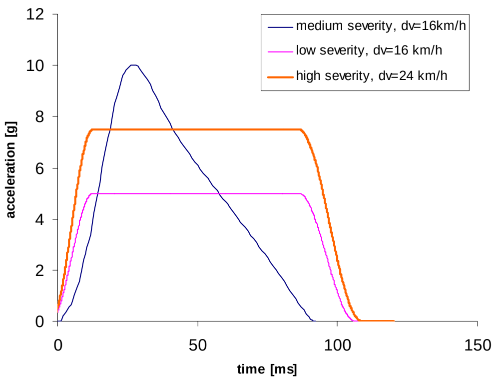

Figure 1. Crash pulses used in the sled tests.

A total of 18 sled tests were conducted using three different crash pulses (Figure 1).

- Pulse 1: low severity pulse, trapezoid, 16 km/h delta-v
- Pulse 2: medium severity pulse, triangular, 16 km/h delta-v
- Pulse 3: high severity, trapezoid-pulse, 24 km/h delta-v

The 18 tests were performed as 6 series whereas each series uses each pulse once (Table 1).

Table 1. Test matix. A total of 18 tests grouped in 6 series were performed.

Test No.	Pulse severity	Series No.
PDB07002	low	1
PDB07001	medium	1
PDB07003	high	1
PDB07004	low	2
PDB07005	medium	2
PDB07006	high	2
PDB07007	low	3
PDB07008	medium	3
PDB07009	high	3
PDB07010	low	4
PDB07011	medium	4
PDB07012	high	4
PDB07017	low	5
PDB07018	medium	5
PDB07019	high	5
PDB07020	low	6
PDB07021	medium	6
PDB07022	high	6

The following measures and neck injury predictors, respectively, were evaluated: NIC [5], Nkm [6], time until dummy head first contacts head restraint (time to head restraint contact), T1-acceleration in x-direction (T1x), rebound velocity, neck shear force (Fx), and neck axial force (Fz).
NIC considers the relative acceleration between head and torso and is derived as shown below.

$$NIC(t) = 0.2m \cdot a_{rel}(t) + (v_{rel}(t))^2 \quad (1)$$

Nkm is calculated by taking into account the neck shear force (Fx) as well as the flexion/extension moment (My).

$$N_{km}(t) = \frac{F_x(t)}{F_{\text{int}}} + \frac{M_y(t)}{M_{\text{int}}} \quad (2)$$

Head restraint contact time was measured by using contact foils. The rebound velocity was derived by film analysis. The maximum rebound-velocity was determined starting at the point in time when the head motion is changing its direction from the rearward to a forward movement. The maximum values of the other criteria were considered from time T0 until the head leaves the head restraint.

Rating system

In general a rating system for assessing the risk of injury for car occupants should be based on biomechanical facts. It should be able to differentiate between high and low injury risk. To obtain usable as well as comparable results it is important to gain robust data with a high level of repeatability and reproducibility. Several rating systems for assessing WAD in low speed rear-impact are already introduced (e.g. IIHS, Thatcham, Folksam, ADAC). EuroNCAP is also planing to implement a new WAD test procedure in their existing occupant rating [4, 7].

This study is based on this proposed EuroNCAP WAD rating system. It consists of two parts, static (geometric) measurements and dynamic sled tests.
For the static tests, the backset and head restraint height were rated according to the limits given in Table 3. Scores range from -1 to +1; a sliding scale was used.
As for the results of the sled tests, Table 4 illustrates the higher and lower performance limits used for the rating. For results in between the higher and lower limits, a sliding scale was used to obtain the score. Each parameter in the dynamic tests can reach a maximum score of 0.5 points. The overall score of a single dynamic test is the sum of the score of NIC, Nkm, rebound velocity, Fx, Fz plus the highest score from either T1x or head contact time i.e. for one pulse a maximum of 3 points is possible.
For the final rating, the worst score of all 3 static measurements (i.e. 3 seats) of one series is added to the points received for all 3 pulses in the dynamic part.

RESULTS

The static evaluation measured by the SAE-Manikin with HRMD of all 18 tests showed a x-value of 45.4 mm (min. 42 mm / max. 52 mm) and a z-value of 37.0 mm (min. 35 mm / max. 39 mm) on average. The BioRID-IIg backset for these tests was 59.0 mm on average (min. 57 mm / max. 60 mm).
The repeatability of the delta-v values reached in the dynamic tests is presented in Table 4. It shows that the delta-v values were generally achieved with high accuracy which results in a low standard deviation.
The results obtained for the parameters determined in the sled tests are summarized in Figures 2 to 9 and Table 6. A repeatability analysis was conducted using the coefficient of variation (CV) method. The CV is defined as the standard deviation (SD) of the measured values divided by the sample mean, and is expressed as a percentage [10]. The repeated responses were assessed by applying the rating scale according Table 2 [11].

Table 2. Rating scale to assess repeatability.

CV = 3%	3% < CV = 7%	7% < CV = 10%	CV > 10
good	acceptable	marginal	not acceptable

It can be seen that T1 acceleration (T1x), head contact time, axial neck force (Fz) and rebound velocity show good to acceptable coefficient of variations (CV). NIC and Nkm, which are derived from basic measures by calculation show a larger CV (marginal to not acceptable). The largest CV, however, was found for the neck shear force (Fx) (not acceptable). This findings corresponds to previous studies [9].

Table 3. Threshold values used for evaluating the static tests.

	Lower performance limit	Higher performance limit
Backset [mm]	30	90
Height [mm]	0	80

Table 4. Preliminary threshold values used for evaluating the dynamic tests [7].

	Lower performance limit	Higher performance limit
Low severity pulse		
NIC [m2/s2]	9	15
Nkm [-]	0.12	0.35
Rebound velocity [m/s]	3.00	4.40
Fx (upper neck shear) [N]	30	110
Fz (neck axial) [N]	270	610
T1 x-acceleration [g]	9.4	12.0
Time head restraint contact [ms]	55	77
Medium severity pulse		
NIC [m2/s2]	11	24
Nkm [-]	0.15	0.55
Rebound velocity [m/s]	3.2	4.8
Fx (upper neck shear) [N]	30	190
Fz (neck axial) [N]	360	750
T1 x-acceleration [g]	9.3	13.1
Time head restraint contact [ms]	51	76
High severity pulse		
NIC [m2/s2]	13	23
Nkm [-]	0.22	0.47
Rebound velocity [m/s]	4.1	5.5
Fx (upper neck shear) [N]	30	210
Fz (neck axial) [N]	470	770
T1 x-acceleration [g]	12.5	15.9
Time head restraint contact [ms]	48	75

Table 5. Delta-v values produced. For each pulse 6 tests were performed.

Pulse	Average delta-v [km/h]	SD [km/h]	CV [%]
1 (low)	15.9	0.3	1.8
2 (medium)	15.7	0.3	2.2
3 (high)	24.1	0.2	0.9

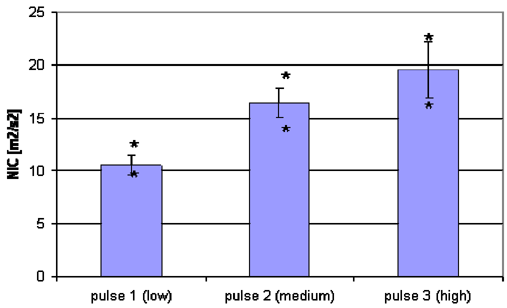

Figure 2. NIC values. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed.

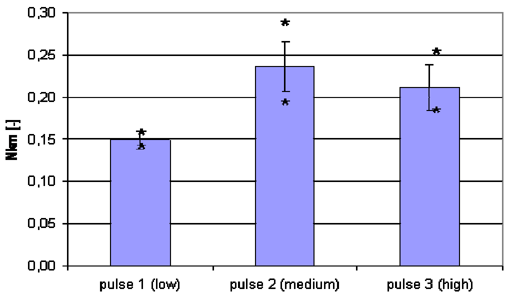

Figure 3. Nkm values. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

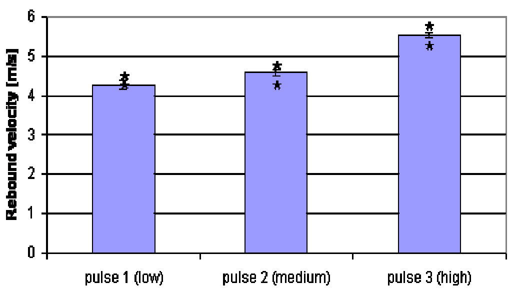

Figure 4. Rebound velocities. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

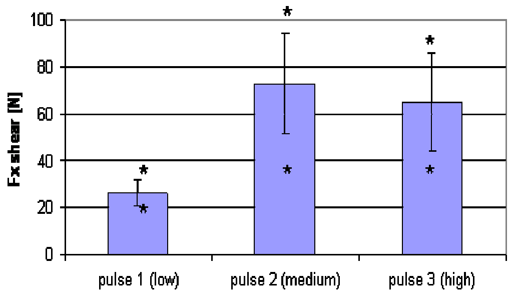

Figure 5. Neck shear forces. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

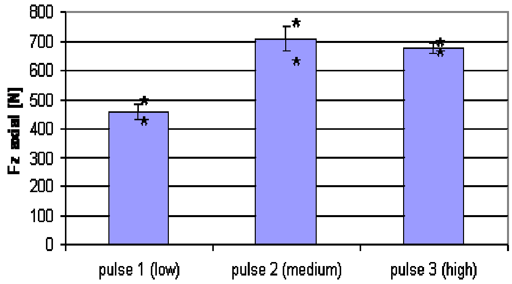

Figure 6. Neck axial forces. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

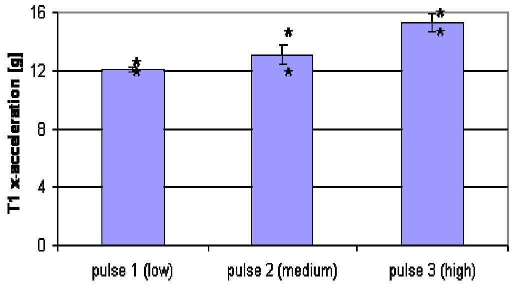

Figure 7. T1 x-acceleration. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

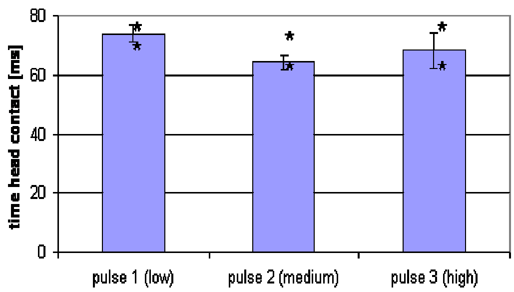

Figure 8. Time to head restraint contact. Error bars denote one standard deviation (SD). Stars represent the minimum and maximum value. For each pulse 6 tests were performed

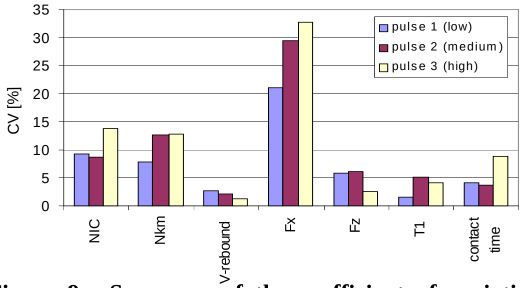

Figure 9. Summary of the coefficient of variation (CV) for all parameters and all pulses.

Table 6. Sled test results. For each pulse 6 tests were conducted.

	Average	Min. Max.	SD	CV [%]
Low severity pulse				
NIC [m2/s2]	10.52	9.73 12.05	0.97	9.18
Nkm [-]	0.15	0.14 0.16	0.01	7.68
Rebound velocity [m/s]	4.28	4.12 4.48	0.12	2.76
Fx (upper neck shear) [N]	26.37	19.17 35.24	5.55	21.04
Fz (neck axial) [N]	456.50	427.00 497.00	26.33	5.77
T1 x-acceleration [g]	12.13	11.82 12.31	0.18	1.45
Time head restraint contact [ms]	73.95	70.0 77.9	2.99	4.05
Medium severity pulse				
NIC [m2/s2]	16.38	14.31 18.52	1.72	10.50
Nkm [-]	0.23	0.19 0.28	0.03	14.02
Rebound velocity [m/s]	4.56	4.45 4.67	0.08	1.83
Fx (upper neck shear) [N]	69.13	36.80 82.30	19.26	27.86
Fz (neck axial) [N]	703.33	622.00 753.00	46.06	6.55
T1 x-acceleration [g]	12.91	11.90 13.65	0.61	4.69
Time head restraint contact [ms]	64.23	61.00 69.00	2.89	4.49
High severity pulse				
NIC [m2/s2]	19.56	15.95 23.36	2.70	13.83
Nkm [-]	0.21	0.18 0.25	0.03	12.73
Rebound velocity [m/s]	5.52	5.44 5.62	0.07	1.23
Fx (upper neck shear) [N]	65.12	35.97 92.30	21.21	32.57
Fz (neck axial) [N]	675.00	644.00 694.00	17.23	2.55
T1 x-acceleration [g]	15.33	14.52 16.10	0.62	4.08
Time head restraint contact [ms]	68.42	62.0 76.0	6.02	8.80

Finally the results were rated according to the scoring system described above. Table 7 summarizes the scores obtained for the static as well as the dynamic tests. The results of the final scores, i.e. adding the worst static and all three dynamic scores for each series, are presented in Table 8.

As an example the final score of series 1 was obtained by adding the scores of the dynamic tests for the 3 different pulses (right column in Table 7) and the worst value of the corresponding static test (middle column in Table 7):

Score dyn. test 1 – low severity pulse	1.80
+ Score dyn. test 1 – medium severity pulse	1.84
+ Score dyn. test 1 – high severity pulse	1.48
+ Score static test 1 – worst value	0.05
Final score series 1	5.17

Table 7. Scores according to the proposed rating system for the static and dynamic tests.

Number of series	Score static test	Score dynamic test
Low severity pulse		
1	0.05	1.80
2	0.10	1.63
3	0.08	1.67
4	0.10	1.48
5	0.05	1.63
6	0.08	1.71
Average	*0.08*	*1.66*
CV %		*6.36*
Medium severity pulse		
1	0.10	1.84
2	0.05	1.62
3	0.05	1.41
4	0.08	1.08
5	0.08	1.33
6	0.08	1.38
Average	*0.07*	*1.44*
CV %		*18.50*
High severity pulse		
1	0.05	1.48
2	0.10	1.44
3	0.10	1.76
4	0.02	1.38
5	0.08	1.24
6	0.10	1.05
Average	*0.08*	*1.39*
CV %		*17.06*

Table 8. Final scores of the 6 test series.

	Final score
series 1	5.17
series 2	4.74
series 3	4.89
series 4	3.96
series 5	4.25
series 6	4.22
Maximum	*5.17*
Minimum	*3.96*
Average	*4.57*
SD	*0.47*
CV [%]	*10.27*

DISCUSSION

In order to investigate the repeatability of the seat assessment procedure according the EuroNCAP proposal [4, 7], a series of sled tests were performed. All tests of this study were conducted with the same seat model whereas for each test a new seat was used. A 3-pulse approach according to the EuroNCAP proposal was applied. Different delta-v values (16 km/h and 24 km/h) as well as different pulse shapes (trapezoid and triangular) were used. This means, that for this kind of WAD assessment of a vehicle seat, a test series of 3 single tests with 3 different pulses is needed. In our study we repeated this complete assessment procedure 6 times, i.e. 18 sled tests were performed (cf. Table 1).

Similarly to the work by Adalian et al. (2005), it could be shown that all crash pulses can be reproduced with sufficient accuracy. The delta-v values for all pulses of this study were achieved with small deviations only.

An important condition was to ensure, that all 18 tests were performed in a very accurate way. Particularly the seat adjustment and the positioning of the BioRID-IIg dummy were set only with less tolerances. The seat back angle was adjusted within a range of 25° +/-0.2° (torso angle of SAE-Manikin). The accuracy in pelvis angle was 26.5° +/-0.3° and the H-Point of the BioRID-IIg was also in a small range of +/-0.3 mm relative to a fixed reference-point. The static measurements (head rest geometry) performed with the HRMD result in a range for the x-value from 42 mm to 52 mm. This was in line with previous studies [9]. The z-values measured by HRMD were within 35 mm to 39 mm. It is well-known, that the backset of the BioRID head has an important influence on the measurements in sled testing. Therefore, we kept this backset as constant as possible. In all of the 18 tests the BioRID-IIg backset was within 57 mm to 60 mm.

All these conditions were important and necessary for this study. To detect changes in the dummy performance, particularly due to tests with the high severity pulse, the dummy was calibrated after the complete test series. Comparing these results with the pre-test calibration did not indicate any changes in the dummy properties.

In a first step the measurements were analysed. In a second step the repeatability of the complete seat assessment method according to the EuroNCAP proposal [4, 7] was investigated.

By comparing the test results between the 3 pulses in detail, it was found that the head-contact time measured in the medium an high severity pulse are almost in the same range whereas the contact time in the low severity pulse is slightly higher on average.

T1x and the rebound velocity show for the low and medium pulse similar values whereas for the high severity pulse the measurements were about 10% to 20% increased.

A significant difference was found in the neck tension force (Fz). The results from medium pulse show the largest deviations (622 N to 753 N) with an average value of 703 N. A surprising result was, that the average value of the high severity pulse (675 N) is lower compared to the medium pulse. Whereas the average in neck tension in low severity pulse tests is remarkable lower (457 N) as expected.

The neck-shear force (Fx) shows the worst result in repeatability. The average values of medium and high severity pulse are almost in the same range (69 N / 65 N). Due to the large deviation of the measurements obtained from the medium and the high severity pulse, the ranges are largely overlapping. This demonstrates that the discriminatory power of such a rating system is limited. The values from low severity pulse clearly indicate a less loading, also with a not negligible deviation from 19 N to 35 N.

The deviation of Nkm, which depends on Fx, shows a less deviation compared to Fx. But due to the overlapping range of measurements (0.19 to 0.28 and 0.18 to 0.25 for medium and high), this criteria is also not able to discriminate between medium and high severity pulse. The average Nkm values for medium and high severity pulse are almost the same, whereas the average value for low severity pulse is lower and shows also a reduced deviation (0.14 to 0.16).

The NIC value on average increases with increasing of loading. (10.5 / 16.4 / 19.6 respectively for low- / medium- / high-severity pulse). Also the deviations increase (CV: 9.2% / 10.5% / 13.8%).

Summarizing it was found, that T1 acceleration (T1x), head contact time, axial neck force (Fz) and rebound velocity show good to acceptable

coefficient of variation [CV]. NIC and Nkm which are derived from basic measures by calculation show a larger standard deviation (marginal to not acceptable). The neck shear force (Fx) showed the largest (not acceptable) spread for all pulses. Depending on the type of pulse (low, medium, high), the differences in CV ranged between 20 % to 30%. Particularly with regard to the very accurate way how the test were prepared and performed this poor repeatability especially for Fx is surprising. An obvious reason could not be found.
Furthermore, the deviation in test results can increase even more by conducting tests in different test labs.

In the next step the single measurements were compared according to the proposed sliding scales [Table 4]. This study investigated only one single seat model. Therefore, an assessment of the sliding scales is not possible. By comparing the measured values with the range of the sliding scale, we assume, that the sliding scale for the head contact time in the low severity pulse is too low compared to medium and high severity pulse and should therefore be moved to higher values. But much more important is the fact, that for most of the criteria the ratio of the range of the sliding scale compared to the range of measured values is questionable, i.e. the deviation of the measured values are too large compared to the sliding scales. The range of measured Fx in medium severity pulse test is 37 N to 92 N, whereas the range for the corresponding sliding scale is 30 N to 190 N [Table 4], that means, the range of measured values spread almost over half of the sliding scale. For NIC and Nkm the spread of the measured values is also not sufficient high compared to the range of the corresponding sliding scales.

Finally we investigated and determined the influence of the measured values on the entire rating scheme [4, 7]
By calculation the rating points for the 3 different pulses we achieved an average of 1.66 points (low severity pulse), 1.44 points (medium severity pulse), 1.39 points (high severity pulse). The coefficient of variation for the low severity pulse is good (CV = 6.36%), whereas for the medium pulse (CV = 18.5%) and the high severity pulse (CV = 17.06%) is not acceptable. The total points for the entire seat assessment including the static measurement are on average 4.57 points with a CV of 10%. Even this deviation does not seem to be remarkable, however, the difference between minimum and maximum score is remarkable. The lowest score of the 6 test series was 3.96 points, the highest 5.17 points; which is about +/-13 % deviation to the calculated mean score, or with other words 26 % from the best to the worst result.

There are legitimate questions if with a rating system which offers such a poor repeatability, an objective seat assessment can be made at all. However, a rating system needs a certain robustness in terms of repeatability otherwise it appears unreliable.
Also there are some concerns about the discriminating power, i.e. to rate a seat against an injury related scale in an objective way.
Each rating system for assessing the occupant safety should be related to a biomechanical scale. But even in the field of WAD this biomechanical knowledge is not complete and therefore derived sliding scales are missed. Nevertheless, at this point we will briefly discuss the biomechanical background. The use of different threshold values for lower and higher performance limits and sliding scales is difficult to understand from a biomechanical point of view. If a criterion is regarded as a predictor for injury, it is assumed that a certain loading results in a corresponding injury risk. Usually biomechanical experiments are the basis on which injury criteria and injury risk curves are defined. In most cases these curves are non-linear and the injury criteria are derived by statistical means (e.g. non-linear regression). Goldsmith and Ommaya (1984), for example, performed several volunteer experiments and derived corresponding threshold diagrams for neck extension/flexion moments as well as for neck shear and axial forces. None of their diagrams shows a linear correlation. Therefore doubts arise with respect to the use of (linear) sliding scales since an evaluation based thereon has hardly any relation to the biomechanical basis of an injury risk.
Similarly, the absolute values chosen in the rating scheme can be criticized. While Goldsmith and Ommaya (1984) found a threshold value for voluntarily tolerated neck shear forces of 845 N the rating system sets an upper limit for the severe pulse of 210 N which is rather low. In contrast, the values for NIC with which a test would be passed go up to 24 m2/s2 in den medium severity pulse. This is not just a value higher than the most often citied injury threshold of 15 m2/s2 but also not logical since the highest values would be expected for the high severity pulse.
Despite the fact that the lower and higher performance limits might lack a biomechanical foundation, adjusting such limits to different crash pulses by means of scaling is fundamentally wrong. From a biomechanical perspective, changing the limits means shifting the threshold on the underlying injury risk curve. In other words, a rating system with different injury threshold values accepts that the occupant is subjected to a different injury risk at a different pulse. Due to the lack of accurate injury risk curves today, the effect of such a shift can not be assessed.

In our paper we criticized the poor repeatability of Fx, at this point we will give an example from the biomechanical perspective. If a person lies on his back on a table such that the head is not supported, an estimated shear force of 48 N (4.8 kg head mass) and a moment of torque of 4.8 Nm (10 cm lever, 4.8 kg) acts on the neck. This already gives a roughly estimated Nkm of 0.15. This opens the question if the proposed sliding scales of these criteria are too low in a region far away where WAD related injuries could occur.

CONCLUSIONS

Performing sled tests representing rear-end collisions revealed that the accuracy with which currently discussed neck injury criteria can be obtained varies between 1.2% and 33%. Since the biomechanical loads discussed in the field of WAD are generally very low in comparison to loads acting in other crash situations, even minor changes in a test set-up may result in significant changes in the loads measured. Consequently, the spread of data increases.

Main Findings

- The neck shear force (Fx) exhibits a "not acceptable" repeatability score for all 3 pulses conducted.
- NIC and Nkm show a "marginal coefficient of variation (CV) in the low severity pulse. However, in the medium and the high severity pulse the CV for NIC and Nkm turn into the "not acceptable" range.
- Although the deviations of most of the single criteria of all 3 pulses are similar. The scoring of the low severity pulse (CV = 6.36%) show less deviation in contrast to the medium (CV = 18.5%) and the high severity pulse (CV = 17.06%).

Rating systems are necessarily based on such test results. Therefore the scoring system used must be robust enough to account for the spread of the input data. Only a comprehensible and repeatable scoring together with a biomechanical relevance will yield to a strong test procedure. The discriminatory power of the scoring system used here, however, seems to be unsatisfactory. The minimum and maximum scores obtained for testing the same seat varied considerably. Consequently, depending on the definition of the final minimum score requirements, the same seat can fail or pass. This finding illustrates a lack of robustness of the scoring system as it is proposed today.

REFERENCES

[1] Adalin C, Sferco R, Ray P (2005) The repeatibility and reproducibility of proposed test procedures and injury criteria for assessing neck injuries in rear impact, Proc. ESV Conf., Paper No. 05-0340

[2] Bortenschlager K, Kramberger D, Barnsteiner K, Hartlieb M, Ferdinand L, Leyer H, Muser M, Schmitt K-U (2003) Comparison tests of BioRID II and RID 2 with regard to repeatability, reproducibility and sensitivity for assessment of car seat protection potential in rear-end impacts, Stapp Car Crash Journal, 47:473-488

[3] RCAR-IIWPG Seat-Head Restraint Evaluation Protocol (Version 2.5 – Sept. 2006)

[4] EuroNCAP-Dynamic Assessment of Car Seats for Neck Injury Protection (Version 2.3 Final Draft)

[5] Boström O, Svensson M, Aldman B, Hansson H, Håland Y, Lövsund P, Seeman T, Suneson A, Säljö A, Örtengren T (1996) A new neck injury criterion candidate based on injury findings in the cervical spinal ganglia after experimental neck extension trauma, Proc. IRCOBI Conf., pp. 123-136

[6] Schmitt K-U, Muser M, Walz F, Niederer P (2002) Nkm - a proposal for a neck protection criterion for low speed rear-end impacts, Traffic Injury Prevention, Vol. 3 (2), pp. 117-126

[7] Klanner W, ADAC, Presentation given on November 14, 2006 at the EuroNCAP Industry Liaison Meeting

[8] Goldsmith W, and Ommaya A. K, (1984) Head and Neck Injury Criteria and Tolerance Levels, The Biomechanics of Impact Trauma. Eds. B. Aldman and A. Chapon, Amsterdam, The Netherlands, Elsevier Science Publishers, 149–187.

[9] Hartlieb M. on behalf of PDB, Study on Repeatability and Reproducibility of BioRID-Dummy Measurements for Whiplash Assessment, ISO/TC22/SC10/WG1 Document-No. N579

[10] Mertz H, (2004), Calculation Methods & Acceptance Levels for Assessing Repeatability and Reproducibility (R & R), ISO/TC22/SC12/WG5 Document-No. N751

[11] Edmund Hautmann, Risa Scherer, Akihiko Akiyama, Martin Page, Lan Xu, Greg Kostyniuk, Minoru Sakurai, Klaus Bortenschlager, Takeshi Harigae, Suzanne Tylko (2003), Updated Biofidelity Rating of the revised WorldSID Prototype Dummy, ESV Conf., Paper No. 388