ASSESSING THE BioRID II REPEATABILITY AND REPRODUCIBILITY BY APPLYING THE OBJECTIVE RATING METHOD (ORM) ON REAR-END SLED TESTS

Linda Eriksson
Autoliv Sverige AB
Sweden
Harald Zellmer
Autoliv B.V. & Co. KG, Elmshorn
Germany
Paper Number 07-0201

ABSTRACT

The BioRID II seems to be the most biofidelic dummy for low-speed rear-end crash tests and is therefore included in several proposed test methods. However, to be broadly accepted, the repeatability and reproducibility of the BioRID II must be verified.

This study aims to assess the BioRID II repeatability and reproducibility by applying the Objective Rating Method (ORM) to rear-end sled tests. The ORM compares crash tests in terms of correlations between criteria, peak values, peak value occurrence times, and curve shapes. Correlations are calculated for all dummy readings and criteria, and for the complete dummy.

Thirty tests were included in this study. These were divided into twelve sets with two to four tests each. The tests within each set mirrored each other, and were used to assess the BioRID II repeatability and reproducibility. The tests were conducted at two crash-test sites. Four BioRID II dummies, five different seats, and three crash pulses were used. Both criteria and dummy readings were compared.

The BioRID II repeatability, in terms of ORM-values, ranged from 83 to 90% with a median value of 88%. Based on component tests with the Hybrid III, TNO/TASS has stated that high correlation is 65% or above. Hence, the BioRID II repeatability is very high. The BioRID II reproducibility ranged from 74 to 78% with a median value of 77%. Five of the nine comparisons included in the reproducibility study were conducted not only with different dummies, but also on different sites.

It can be concluded that the BioRID II shows high repeatability and reproducibility for all of the compared crash conditions. Furthermore, the BioRID II shows excellent repeatability for nearly all of the NIC, N_{km}, T1x, HC, F_x, and M_y criteria comparisons. The ORM-values for these criteria were predominantly above 90%.

INTRODUCTION

The BioRID II seems to be the most biofidelic dummy for low-speed rear-end crash tests ([1], [2], [3], [4], [5]) and it has been shown that the BioRID II is a good tool for predicting low-speed rear-end neck injuries ([6], [7], [8]). Therefore, the BioRID II is included several proposed test methods, among them the EuroNCAP Final Draft [9].

Several studies have been performed to evaluate the BioRID II repeatability. Good repeatability was shown already for early versions of the BioRID by [11] and [12]. [13] evaluated BioRID II (however not the actual version g) repeatability by exposing one dummy seated in four different seat designs three times to a 16 km/h crash pulse. Also the reproducibility was evaluated by using three different dummies on a rigid steel seat. The BioRID II showed sufficiently good repeatability and reproducibility, although these were somewhat better for the RID2 which was also included in the study.

[14] performed three repeated tests on three different seats using a 16 km/h crash pulse. To evaluate reproducibility, the same three seats were tested at five different test labs using two different crash pulses (16 and 25 km/h). The sleds used included both acceleration and deceleration types. Scattering was defined as difference between maximum and minimum values divided by the mean value. Repeatability was rated to be good – meaning scattering being about 20% – and reproducibility was rated acceptable at 16 km/h (scattering 10% to 40%). But the scattering at 25 km/h showed to be generally between 30% and more than 100% on biomechanical criteria. The authors mentioned that training in seat and dummy set-up will help to improve the results.

[15] carried out repeatability and reproducibility investigations using four BioRID dummies, two types of seats, and two crash pulses (16 and 24 km/h). All together thirty-eight tests were performed at one test facility. Almost the same NIC values were found for all four dummies. The repeatability

(maximum deviation from the mean value) for NIC was ±13%, the reproducibility (maximum deviation from the mean value of all four dummies) for NIC was ±3.5%. The repeatability for N_{km} was ±20% and the reproducibility ±11%. The repeatability for F_x was ±34 %, and the reproduceibility was ±27 % with clear dependency on dummy used. The repeatability for F_z was ±6 % and the reproducibility was ±8 %. The influence of pulse variation inside its corridors was studied but did not show apparent influence on the before mentioned measurement values.

[16] investigated repeatability and reproducibility of dummy response using the coefficient of variation (CV). Three BioRID dummies underwent five tests on a rigid seat design each. The CV for repeatability was expressed as percentage after dividing the standard deviation of the peak measurement values for each dummy by the average value. The CV for reproducibility was also expressed as percentage after dividing the standard deviation of differences among the three dummies by the average value. Repeatability in terms of CV showed to be below 5% for head acceleration and neck moment in flexion, between 5% and 10% for F_x, F_z, and T1 acceleration, and in some cases slightly above 10% in neck extension. Reproducibility in terms of CV was 6.3% for T1 acceleration, 5.0% for F_x, 13.7% for F_z, 3.3% for neck flexion, and 31.6% for neck extension.

METHOD

Twelve sets of rear-end sled tests were evaluated in this study. Each set of crashes contained two to four tests, and all tests within each set were designed to mirror each other. Altogether, thirty tests were included in the twelve sets. Tests with four different BioRID II dummies (all of version g), three different crash pulses, five different seats, and conducted at two different crash sites, were included in order to assess the BioRID II repeatability and reproduce-ibility. The crashes were conducted between November 2004 and August 2006. The test set-ups can be found in Table 1.

The seats used were standard seats with head restraints. However, some of them had added, excluded, or modified safety systems, or were tested during development. The three crash pulses used were those that are proposed for EuroNCAP, how-ever all of them do not fulfil all pulse requirements specified in the EuroNCAP v2.3 Final Draft ([9]). The IIWPG pulse is triangular shaped with a peak of appr. 10g at 27 ms, mean acceleration of 5.5g, and Δv of 16 km/h. The Folksam/SRA low severity pulse (FV1) is trapezoidal shaped with mean acceleration of 4g and Δv of 16 km/h, and the Folksam/SRA high severity pulse (FV2) is trapezoidal shaped with mean acceleration of 6.5g and Δv of 24 km/h. The two crash sites used were those located at Autoliv in Vårgårda, Sweden (ALS) and at Autoliv in Elmshorn, Germany (ANG). At ALS a hydraulic accelerator sled was used, and at ANG a deceleration sled with a hydraulic brake system was used.

Seventeen pairs of tests were conducted with exactly the same set-up and these were used to assess the BioRID II repeatability. Set 4 contained four tests with two different dummies, while all other variables were similar. These tests were used to assess the BioRID II reproducibility. Set 11 contained two tests conducted at different tests sites and with different dummies and showed a combined correlation for dummy reproducibility and test repeatability at different sites. Set 12 also contained two pairs of tests conducted at different sites and with different dummies, and were used to assess both repeatability and reproducibility.

Table 1.
Included tests

Set	Test	Dummy	Seat	Pulse	Site	Test Date
1	**a**	B1	A	IIWPG	ALS	Dec. 2004
	b	B1	A	IIWPG	ALS	Dec. 2004
2	**a**	B1	B	FV3	ALS	April 2005
	b	B1	B	FV3	ALS	April 2005
	c	B1	B	FV3	ALS	April 2005
3	**a**	B1	A	FV3	ALS	April 2005
	b	B1	A	FV3	ALS	April 2005
	c	B1	A	FV3	ALS	April 2005
4	**a**	B2	A	IIWPG	ALS	Feb. 2006
	b	B2	A	IIWPG	ALS	Feb. 2006
	c	B1	A	IIWPG	ALS	Feb. 2006
	d	B1	A	IIWPG	ALS	Feb. 2006
5	**a**	B1	A	IIWPG	ALS	Sept. 2005
	b	B1	A	IIWPG	ALS	Sept. 2005
6	**a**	B1	A	FV1	ALS	Sept. 2005
	b	B1	A	FV1	ALS	Sept. 2005
7	**a**	B1	A	FV3	ALS	Nov. 2005
	b	B1	A	FV3	ALS	Nov. 2005
8	**a**	B3	C	FV3	ANG	Aug. 2006
	b	B3	C	FV3	ANG	Aug. 2006
9	**a**	B3	C	FV1	ANG	Aug. 2006
	b	B3	C	FV1	ANG	Aug. 2006
10	**a**	B3	C	IIWPG	ANG	Aug. 2006
	b	B3	C	IIWPG	ANG	Aug. 2006
11	**a**	B1	D	IIWPG	ALS	Nov. 2004
	b	B4	D	IIWPG	ANG	Nov. 2004
12	**a**	B1	E	IIWPG	ALS	Nov. 2004
	b	B1	E	IIWPG	ALS	Nov. 2004
	c	B4	E	IIWPG	ANG	Nov. 2004
	d	B4	E	IIWPG	ANG	Nov. 2004

Within each set, the same dummy positioning procedure was used. The compared dummy records were the x- and z-accelerations in the head, C4, T1, T8, L1 and pelvis, and the upper neck shear force (F_x), tension (F_z) and bending moment (M_y). The filter classes used were those specified in the Euro-NCAP v2.3 Final Draft ([9]): CFC 1000 for head z-acceleration, F_x and F_z, CFC 600 for M_y, and CFC 60 for all other. However, for test 10b, the C4, T1, T8, L1, and pelvis accelerations were filtered with CFC 180 and the head x-acceleration with CFC 1000. In test 10b the T1 z-acceleration was missing, and in tests 11b and 12c the T8 z-accelerations were missing. Further, the crash pulses (filtered with CFC 60) were compared for all sets with the exception of set 9, since the crash pulse in test 9a was missing. The NIC, the N_{km}, the maximum T1 x-acceleration, the head restraint contact time, the maximum upper neck shear force, and the maximum upper neck tension force were calculated for all tests according to the EuroNCAP v2.3 Final Draft ([9]).

The Objective Rating Method, ORM, ([10]) was used to assess the correlation between the tests in each set. The ORM correlates one comparison test to one reference test. Therefore, the ORM was applied to all possible pairs in those sets that contained more than two tests. The ORM enables comparison between scalars, such as criteria, minimum and maximum peak values, and their occurrence times, and between curve shapes.

In this study the criteria and signals listed in Table 2 were included. The acceleration and the neck load maximum peak values and their occurrence times were compared, and for the M_y also the minimum peak value and its occurrence time were compared. The peak value comparisons were limited to peaks occurring during the first 200 ms of the crashes. Further, the curve shapes of the signals were compared during the first 200 ms of the crashes, and the crash pulse shapes were compared during the first 100 ms of the crashes.

Table 2.
Compared criteria and signals

Group	Component
Criteria	NIC (max. before end of head contact) N_{km} (max. before end of head contact) T1x (max. before end of head contact) HC (head restraint contact time) F_x (max. before end of head contact) F_z (max before end of head contact)
Acceleration	Head x-acc Head z-acc C4 x-acc C4 z-acc T1 x-acc T1 z-acc T8 x-acc T8 z-acc L1 x-acc L1 z-acc Pelvis x-acc Pelvis z-acc
Neck Loads	F_x F_z M_y
Crash pulse	Crash pulse

The ORM scalar correlations are calculated according to Equation 1. This expression is called the Factor Method and calculates the correlation between the reference test and the comparison test. The results range from 0 to 100%, where 100% represents a perfect match.

The curve shape correlation is calculated according to Equation 2. This expression is called the Weighted Integrated Factor Method and is a combination of the Factor Method and the Root Mean Square Addition Method. This means that the correlation in each time step contributes to the total correlation just as the function value would contribute to the total area underneath the curve. The δ is very small and used to avoid division by zero. *r* and *c* are used as abbreviations for *reference* and *comparison*, respectively.

$$ORM_{scalar} = \frac{\max(0, reference \cdot comparison)}{\max(reference^2, comparison^2)} \quad (1)$$

$$ORM_{shape} = 1 - \sqrt{\frac{\int \left(\max(|r(t)|, |c(t)|) \cdot \left(1 - \frac{\max(0, r(t) \cdot c(t))}{\max(\delta, r(t)^2, c(t)^2)} \right)^2 \right) dt}{\int \max(|r(t)|, |c(t)|) dt}} \quad (2)$$

In order to simplify comparison between tests, ORM-values are calculated not only for the scalars and the curve shapes, but also for groups of scalars and curve shapes. The contribution of each ORM-value in its group is defined by a weight factor (W). Equation 3 is used to calculate the ORM-value for each group. Further, the groups are arranged into one single ORM-value that is the correlation for the complete system. The contribution of the group ORM-values to the complete ORM-value is defined by weight factors (W). Equation 4 is used to calculate the ORM-value for the complete system.

$$ORM_{group} = 1 - \sqrt{\frac{\sum \left(W_{scalar\ or\ shape} \cdot (1 - ORM_{scalar\ or\ shape})^2 \right)}{\sum W_{scalar\ or\ shape}}} \quad (3)$$

$$ORM_{complete} = 1 - \sqrt{\frac{\sum \left(W_{group} \cdot (1 - ORM_{group})^2 \right)}{\sum W_{group}}} \quad (4)$$

In this study, the criteria were collected in the group Criteria, the head, the spine, and the pelvis accelerations were collected in the group Acc, and the neck loads were collected in the group Neck Loads. The weight factors (W) used for the scalars and curve shape to form the ORM-values for the groups are listed in Table 3. The weight factors used for the complete ORM-value were 6 for the Criteria since this group included six criteria, 12 for the Acc since this group included signals from six accelerometers in two directions (x and z), and 3 for the Neck Loads since this group included three signals from the upper neck load cell (see Table 3). For tests with one missing acceleration signal (10b, 11b, and 12c), the weight factor used for the group Acc was 11. The tests in the set 9 and 10 did not have any positive F_x peaks, and consequently no ORM-values for the maximum peak or its occurrence time could be used. Therefore, the weight factor for the group Neck Loads was reduced by 2/3 to 2.333 for set 9 and 10. The crash pulses were not included in the complete ORM-values, since these were aimed to show the BioRID II repeatability or reproducibility.

Table 3
Weight factors used

Group	W_{group}	Component		$W_{scalar\ or\ shape}$
Criteria	6	NIC		0.167
		N_{km}		0.167
		T1x		0.167
		HC		0.167
		F_x		0.167
		F_z		0.167
Acc	12	Head x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		Head z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		C4 x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		C4 z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		T1 x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		T1 z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		T8 x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		T8 z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		L1 x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		L1 z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		Pelvis x-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		Pelvis z-acc	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
Neck Loads	3	F_x	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		F_z	Max peak	0.028
			Max peak time	0.028
			Curve shape	0.028
		M_y	Max peak	0.014
			Max peak time	0.014
			Min peak	0.014
			Min peak time	0.014
			Curve shape	0.028

RESULTS

BioRID II Repeatability

Seventeen pairs of tests were used to assess the BioRID II repeatability. The ORM-values for the complete system correlation range from 83 to 90% (median value 88%) and are shown in Figure 1. The ORM-values of the group Criteria for the same sets range from 89 to 97% and are shown in Figure 2, and their components are listed in Table 4. Among the Criteria components the F_x shows the largest spread, from 78 to 100%, and the lowest median value of 91%. The other five criteria have median values of 95% or above.

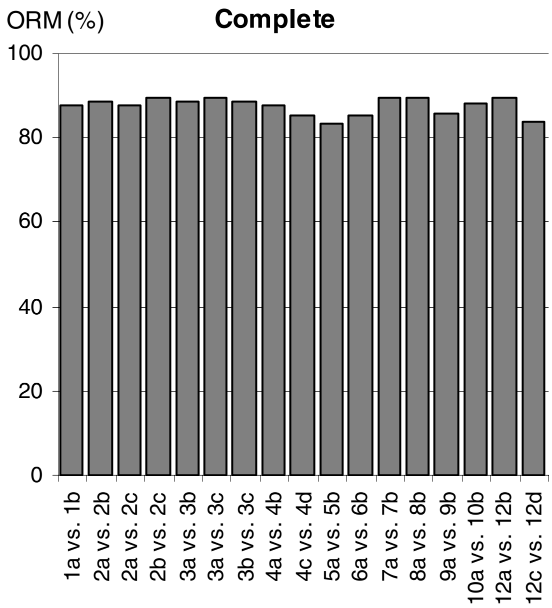

Figure 1. BioRID II repeatability tests: ORM-values in percent for the complete system

The ORM-values of the group Acc range from 82 to 90% and are shown in Figure 3. The ORM-values for the group Neck Loads range from 77 to 92 and are shown in Figure 4. In general, the peak value correlations are high; the median ORM-values for all peaks are shown in Figure 5. The ORM-values for the peak occurrence times are in general very high; the median values are shown in Figure 6. In total, the peaks for 268 pairs were compared. Of these, 23 match perfectly and only one peak ORM-value was below 65%. Among the peak value occurrence times 45 pairs match perfectly and three pairs have ORM-values less than 65%. Two of these three pairs have double peaks of almost the same magnitude in the T8 z-acceleration signals that cause the low correlation values: in both reference tests the first peak is the highest one and in the both comparison test the latter peak is the highest. For the third pair, there are double peaks in the M_y signals.

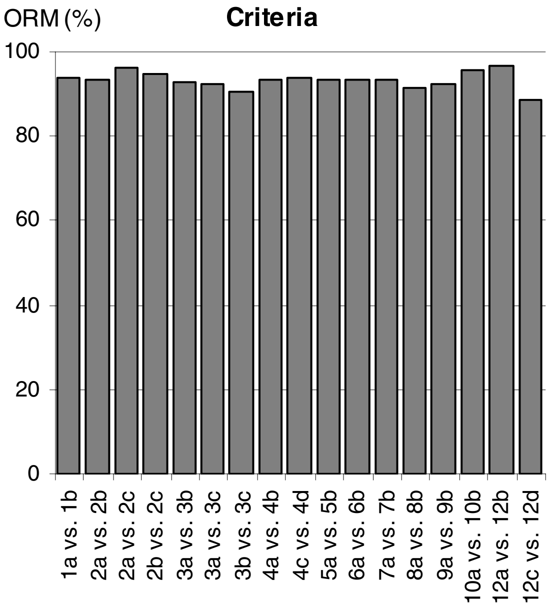

Figure 2. BioRID II repeatability tests: ORM-values in percent for the group Criteria.

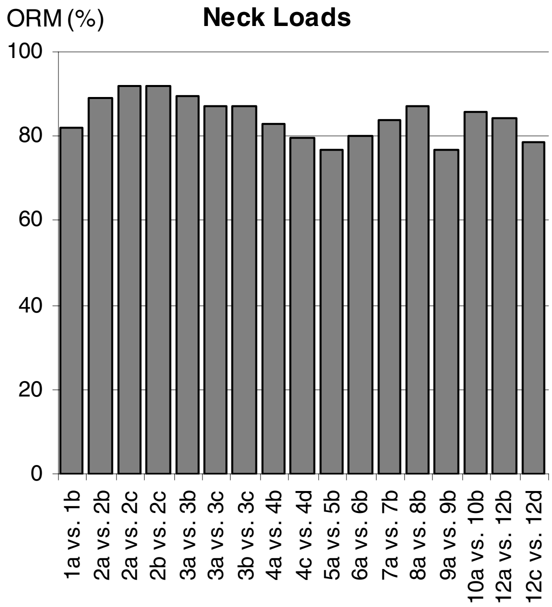

Figure 4. BioRID II repeatability tests: ORM-values in percent for the group Neck Loads.

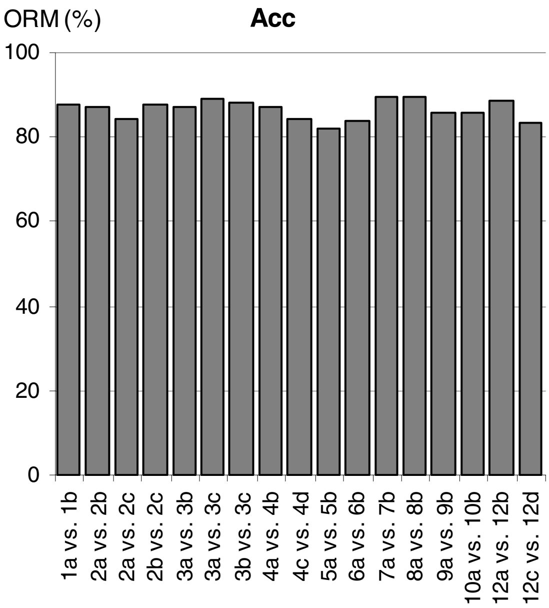

Figure 3. BioRID II repeatability tests: ORM-values in percent for the group Acc.

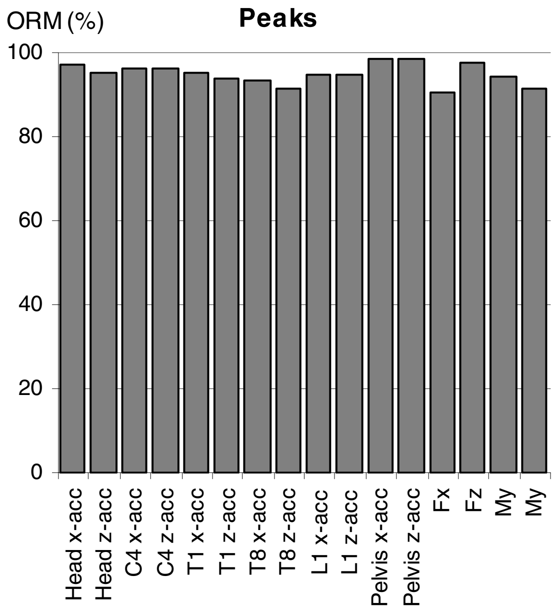

Figure 5. BioRID II repeatability tests: median ORM-values in percent for the maximum acceleration and neck load peak values, the last bar is the median minimum M_y peak value.

Table 4
BioRID II repeatability tests: ORM-values in percent for the components in the group Criteria

Tests	NIC	N_{km}	T1x	HC	F_x	F_y
1a vs. 1b	99	92	88	99	96	99
2a vs. 2b	92	92	95	96	90	98
2a vs. 2c	99	95	93	98	100	97
2b vs. 2c	93	97	98	98	91	99
3a vs. 3b	87	100	94	98	91	95
3a vs. 3c	91	97	98	100	85	94
3b vs. 3c	96	97	96	98	78	98
4a vs. 4b	99	91	95	100	88	96
4c vs. 4d	98	100	92	98	87	98
5a vs. 5b	96	86	95	97	95	96
6a vs. 6b	99	89	97	89	94	99
7a vs. 7b	90	96	89	100	96	99
8a vs. 8b	98	92	97	100	82	95
9a vs. 9b	86	95	98	95	100	90
10a vs. 10b	93	95	98	94	100	97
12a vs. 12b	99	93	98	99	97	98
12c vs. 12d	79	92	93	97	85	99
Min value	79	86	88	89	78	90
Median value	96	95	95	98	91	98
Max value	99	100	98	100	100	99

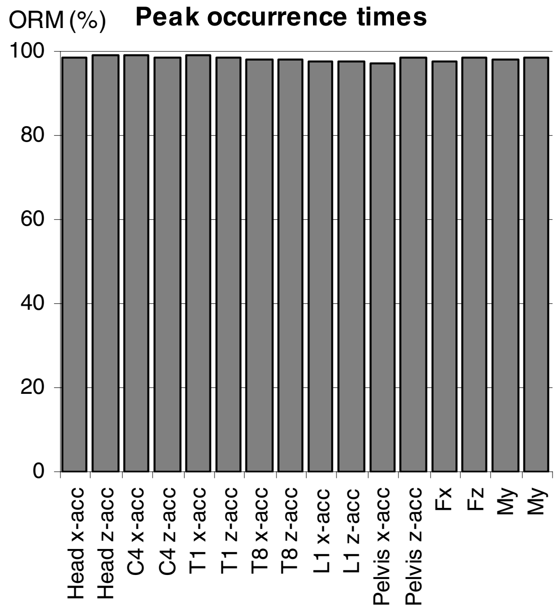

Figure 6. BioRID II repeatability tests: median ORM-values in percent for the occurrence time for the maximum acceleration and neck load peak values, the last bar is the median occurrence time for the minimum M_y peak value.

Figure 7 shows the ORM-values for the median curve shape correlations. The lowest median ORM-values can be found for T1 z-acceleration, T8 z-accelerations, and M_y. These, together with F_x, are the only signals for which the lowest curve shape ORM-values are below 65%. Still, all median ORM-values are above 65%.

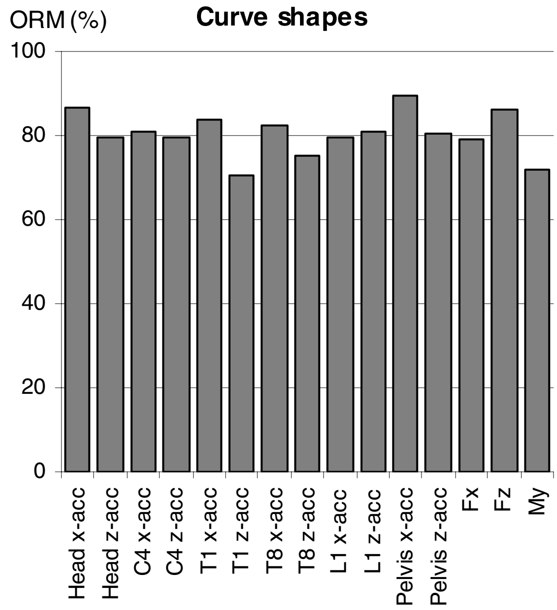

Figure 7. BioRID II repeatability tests: median ORM-values in percent for the acceleration and neck load curve shapes.

For the seventeen pairs of tests used to assess the BioRID II repeatability the corresponding ORM-values for the crash pulse shapes are shown Figure 8.

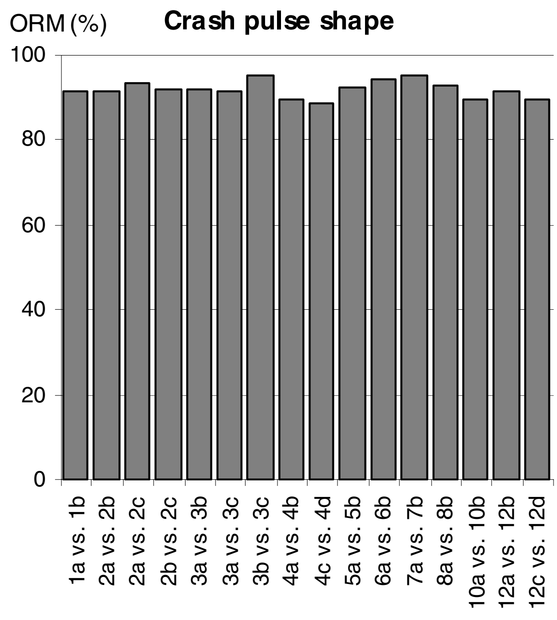

Figure 8. ORM-values in percent for the crash pulse shapes for the tests included in the BioRID II repeatability assessment. There is no ORM-value for set 9 because of a missing crash pulse.

BioRID II Reproducibility

Three sets can be used to assess the BioRID II reproducibility: Set 4, Set 11, and Set 12. In Figure 9 the ORM-values for the crash pulse shapes are given. The correlations of the crash pulse shapes are high for all pairs in Set 4 but somewhat lower for the reproducibility tests in Set 11 and Set 12. Set 4 was designed to evaluate the repeatability and these tests were conducted at ALS. In Set 11 and Set 12, not only two dummies were used, also two crash sites were used.

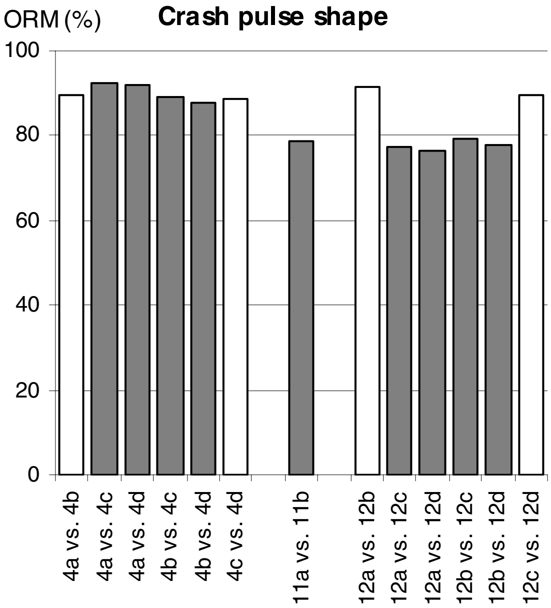

Figure 9. ORM-values in percent for the crash pulse shape. White bars are used for the BioRID II repeatability tests, and grey bars are used for the BioRID II reproducibility tests.

The ORM-values for the complete BioRID II reproducibility are shown in Figure 10. The reproducibility are lower than the repeatability, nevertheless well above 65%. The BioRID II reproducibility ORM-values range from 74 to 78%, with a median value of 77%. This should be compared with the range from 83 to 90% (median value 88%) for the BioRID II repeatability (Figure 1).

The ORM-values for the groups are shown in Figure 11 (Criteria), Figure 12 (Acc), and Figure 13 (Neck Loads). As can be seen, for the groups only one ORM-value is below 65%. That is the group Neck Loads for test 4b versus 4d that has a ORM-value of 62%, mainly because of low curve shape ORM-values for F_x and M_y. The ORM-values for the components in the groups Criteria are given in Table 5. The F_x is below 65% for there of the nine cases, the median value for this criteria is 68%. The corresponding value was 91% for the BioRID II repeatability tests. The other five criteria shows much better correlations than the F_x do, however the median values for all criteria are less good than for the BioRID II repeatability tests. The differences for these five criteria are between 2 and 9 percent units.

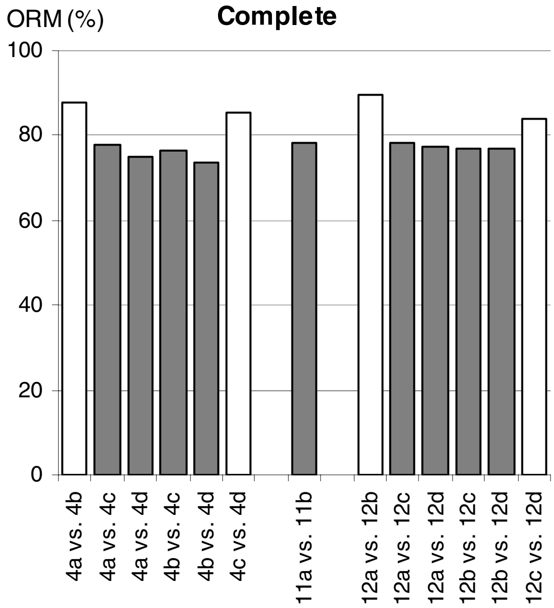

Figure 10. ORM-values in percent for the complete systems. White bars are used for the BioRID II repeatability tests, and grey bars are used for the BioRID II reproducibility tests.

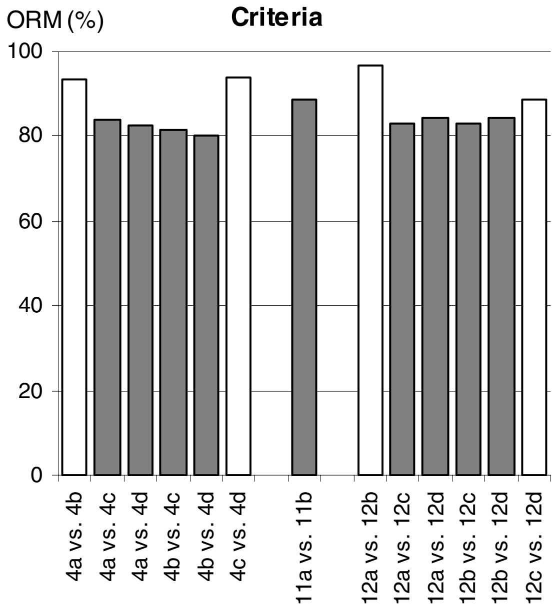

Figure 11. ORM-values in percent for the groups Criteria. White bars are used for the BioRID II repeatability tests, and grey bars are used for the BioRID II reproducibility tests.

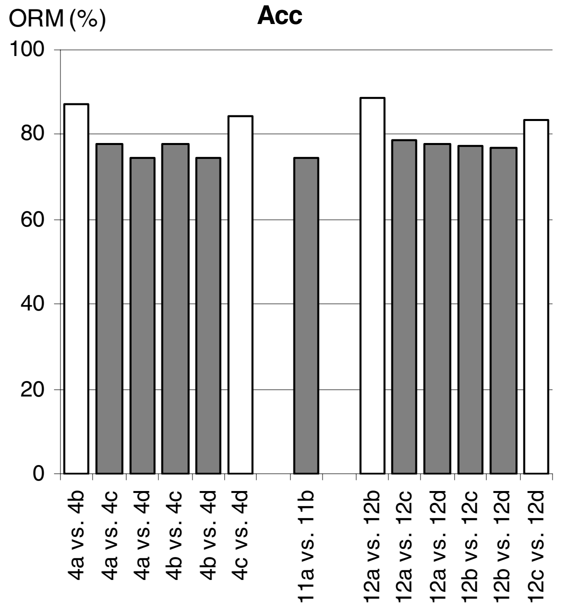

Figure 12. ORM-values in percent for the groups Acc. White bars are used for the BioRID II repeatability tests, and grey bars are used for the BioRID II reproducibility tests.

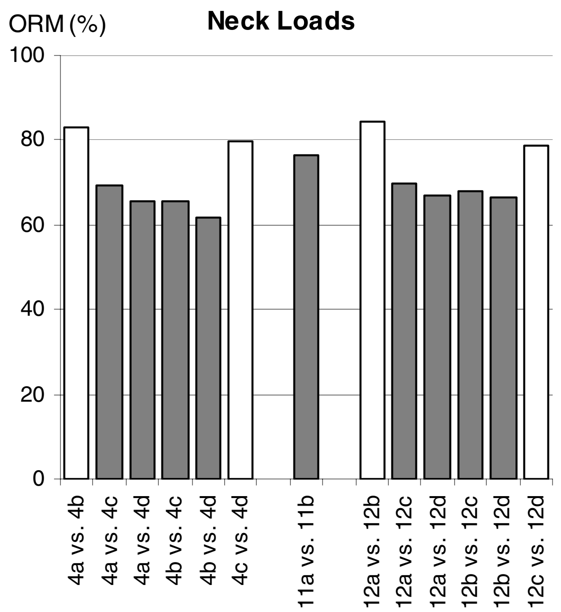

Figure 13. ORM-values in percent for the groups Neck Loads. White bars are used for the BioRID II repeatability tests, and grey bars are used for the BioRID II reproducibility tests.

Table 5.
BioRID II reproducibility tests: ORM-values in percent for the components in the group Criteria

Tests	NIC	N_{km}	T1x	HC	F_x	F_y
4a vs. 4c	88	83	80	98	77	87
4a vs. 4d	86	83	88	95	67	89
4b vs. 4c	89	92	76	98	68	84
4b vs. 4d	87	92	84	95	59	86
11a vs. 11b	87	78	95	98	96	92
12a vs. 12c	96	89	86	96	64	91
12a vs. 12d	82	82	92	93	75	90
12b vs. 12c	96	89	86	96	64	91
12b vs. 12d	82	88	90	94	73	88
Min value	82	78	76	93	59	84
Median value	87	88	86	96	68	89
Max value	96	92	95	98	96	92

The median ORM-values for the peak values in the BioRID II reproducibility tests are shown in Figure 14. The median ORM-values for the peak values are considerably lower for the T1 z-acceleration and the F_x compared to the other. However, taking the ranges into account, also the M_y values are low. Three cases match perfectly for the peak values. The numbers for the peak value occurrence times are somewhat better, for fifteen pairs the peak time correlated with 100%. The median ORM-values for peak occurrence times are given in Figure 14. These are nearly as high as those for the BioRID II repeatability tests (Figure 6). Only the median ORM-values for the T1 z-acceleration and the M_y negative peak occurrence times are significantly lower.

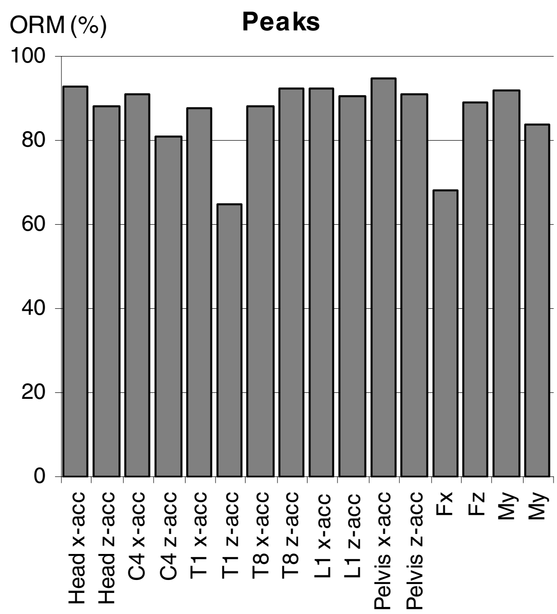

Figure 14. BioRID II repoducibility tests: median ORM-values in percent for the maximum acceleration and neck load peak values, the last bar is the median minimum M_y peak value.

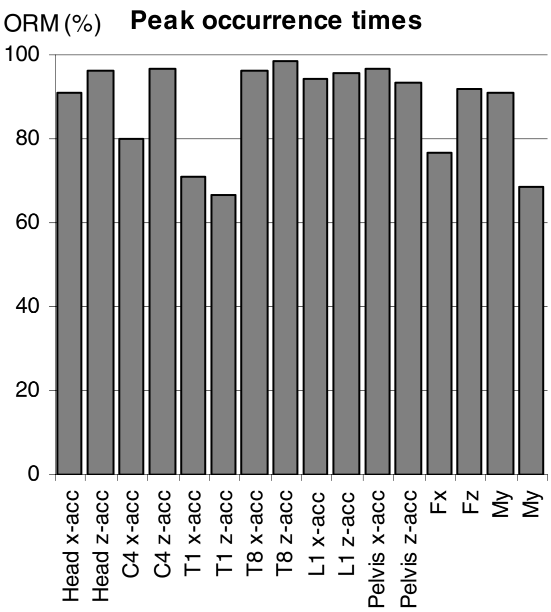

Figure 15. BioRID II reproducibility tests: median ORM-values in percent for the maximum acceleration and neck load peak value occurrence times, the last bar is the median minimum M_y peak value occurrence time.

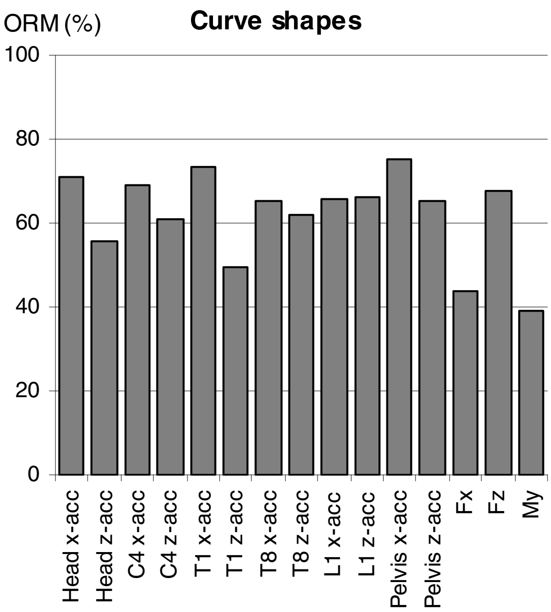

Figure 16. BioRID II reproducibility tests: median ORM-values in percent for the acceleration and neck load curve shapes.

The median ORM-values for the curve shapes are much lower for the reproducibility tests compared to the repeatability tests. The z-accelerations for the head, C4, T1, and T8, and the F_x and M_y median curve shape ORM-values are below 65%; for the repeatability tests all were above. In general the curve shape trends are similar for all compared pairs, but the magnitudes for the peaks differ and that results in somewhat lower ORM-values. Predominantly, the least correlating parts of the curves occur between 150 ms and 200 ms.

DISCUSSION

Repeatability and reproducibility studies

The values for repeatability and reproducibility from the studies conducted by [11], [12], [13], [14], [15], [16] cannot be directly compared to each other or to this study since each study calculate repeatability and reproducibility with different methods. [14] assessed good repeatability in general, acceptable reproducibility at 16 km/h, and unacceptable reproducibility at 25 km/h. It has to be mentioned that all biomechanical values were largely exceeding common thresholds in the 25 km/h tests. [15] and [16] both did tests resulting in low to medium bio-mechanical values but showing just opposite trends in F_x and F_z reproducibility. In this study, good overall repeatability and reproducibility were assessed for the BioRID II. However, the F_x was below the limit for good reproducibility in three of nine comparisons, but it showed good repeatability for all seventeen comparisons.

High Repeatability: ORM > 65%

The Objective Rating Method (ORM) was published in 2005 ([10]) as a tool for assessing the correlation of Madymo simulation models to mechanical tests. They stated that high correlation is 65% or above repeatability for mechanical tests. This statement was based on component tests on one Hybrid III 50%-ile without arms and lower legs. Ten different tests were repeated ten times, and in each test thirty signals were recorded. All signals of the repeated tests were then compared to the first test in each test series. However, which signals that were compared, or their weight factors, are not specified. According to the authors, special attention was taken in positing the dummy before each test to ensure good repeat-ability and a well-defined environment was used in the tests. In this study, the ORM-values for the BioRID II repeatability ranged between 83 and 90% with a median value of 88%. This is much better numbers than those presented by [10] for Hybrid III component tests. Take into consideration that for the BioRID II tests, not only the BioRID II spread are measured, also the spread in the seats and test environments are included. Hence, it can be assessed that the BioRID II repeatability is very high.

The BioRID II repeatability assessment was based on seventeen pairs of test. In these tests four different BioRID II dummies, four seats, three pulses, and two sites were used (Table 1). The specific influences of these parameters on the ORM-values cannot be assessed since seventeen pairs are too few to measure significant differences. Further, the spread is between 83 and 90%, and twelve of the seventeen pairs have ORM-values of 88, 89 or 90%. Consequently, there are only minor differences between most of the pairs in terms of repeatability ORM-values. Although no significant influences can be assessed, the BioRID II repeatability ORM-values are presented sorted according to dummies (Figure 17), seats (Figure 18), pulses (Figure 19), and sites (Figure 20) used. The unsorted data can be found in Figure 1. The B2 and B4 BioRID II dummies were only used for one test each. Therefore, no conclusions should be draw regarding these dummies. The median ORM-values for B1 and B3 are almost the same. The spread when using the same dummy appear to be somewhat wider than the spread between different dummies (Figure 17). The spread in ORM-values do not likely depend on the seat (Figure 18) or the site (Figure 20). Further, the spread is much smaller between the tests conducted with FV3 pulse than for the other pulses (Figure 19).

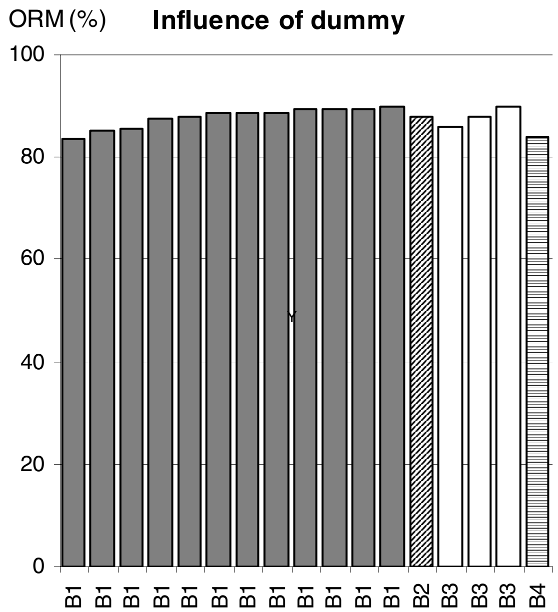

Figure 17. BioRID II repeatability tests: ORM-values for the complete system sorted according to seat used.

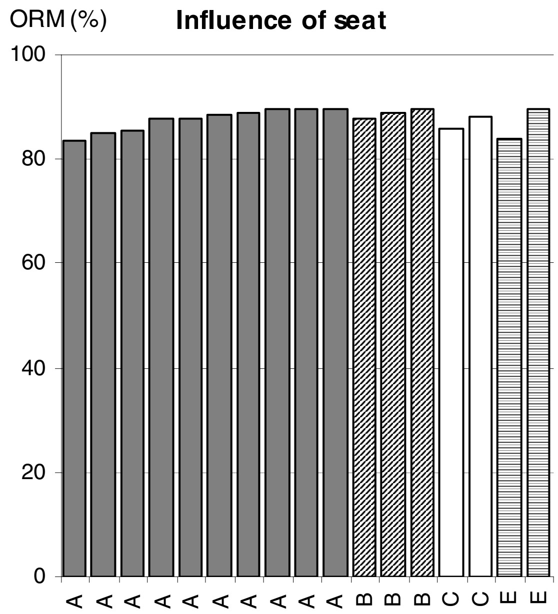

Figure 18. BioRID II repeatability tests: ORM-values for the complete system sorted according to seat used.

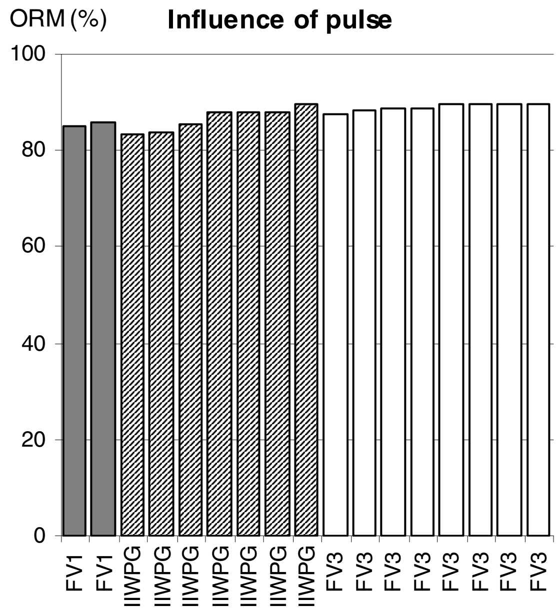

Figure 19. BioRID II repeatability tests: ORM-values for the complete system sorted according to pulse used.

Set 2 and Set 3 contained three tests each that were set-up to mirror each other. The NIC, the N_{km}, the T1x, the HC, the F_x, and the M_y values for these sets were analyzed in order to find out if these values increased or decreased with the number of tests. Only the T1x values for Set 2 shows continues decrease, and none shows continues increase. Hence, nothing in this study indicates that the BioRID II responses changes due to the number of conducted tests.

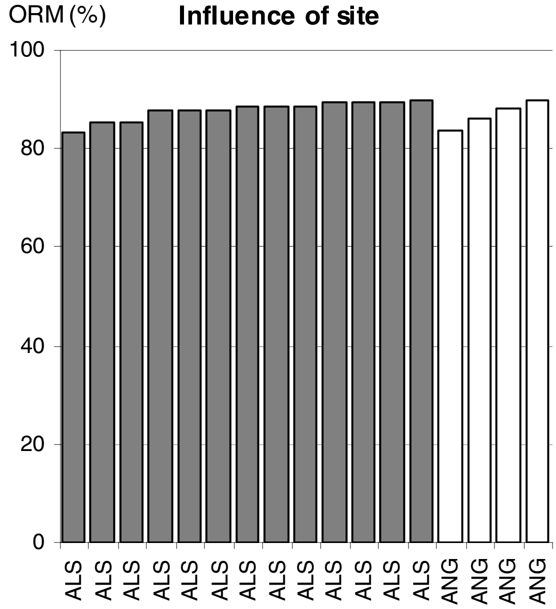

Figure 20. BioRID II repeatability tests: ORM-values for the complete system sorted according to site used.

Three sets were used to assess the BioRID II reproducibility. Of these, only Set 4 were designed with the aim to evaluate the reproducibility. Set 11 and Set 12 were parts of a seat improvement study that were conducted at two sites. Therefore, not only the dummies used differed, also the crash sites differed. As can be seen in Figure 9, the crash pulses shapes were less similar for the reproducibility tests conducted at two different sites (Set 11 and 12) than those conducted at the same site (Set 4). A comparison between the pulses used in Set 11 is given in Figure 21. Likely, the differences between these crash pulses only influence the outcome negligible. However, there are other differences that may have influenced the outcome. The gap between the dummy and the head restraint differed between the sites. Therefore, it can not be excluded that the dummy positions influenced the outcome. For Set 11, the gap was 5 mm wider for the dummy at ALS, and for Set 12 the gaps were 7 and 9 mm wider at ALS. Furthermore, the crash tracks used at ALS and ANG differ. At ALS a hydraulic acceleration sled is used: the dummy is at rest when the crash starts and by the aid of a hydraulic system the dummy is accelerated with a pre-defined pulse. At ANG a Hydro-Brake sled is used: prior to the crash the dummy is moving and a hydraulic system is then used to brake the sled with a pre-defined deceleration pulse. According to Figure 10 the ORM-values for the BioRID II reproducibility tests are in the same range for all three sets. Hence, it is likely that the dummy influence in much larger than influence from the positioning and the test conditions.

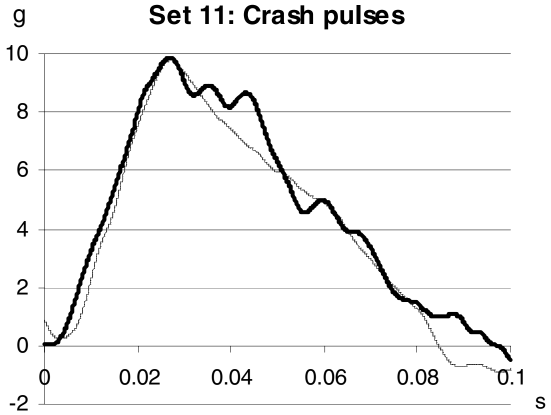

Figure 21. The crash pulses used in Set 11. The reference test (thick line) was conducted at ALS and the comparison test (thin line) at ANG.

The ORM-values for the criteria values, the peak values, and the peak value occurrence times are generally higher than ORM-values for curve shapes. However, both equation 1 and 2 will result in the same ORM-value if they are applied to the same scalars. Nevertheless, our demands are often much higher on scalars than on curves. Hence, different rule of thumbs may be used when deciding if scalars or curves correlate well. Up to date, too few correlation evaluations have been performed to assess if different type of tests and measured signals require different ORM threshold values.

The ORM

An example from Set 4 will be presented in detail in order to provide a feeling for the ORM scale. This set contains four tests: 4a and 4b were conducted with one BioRID II dummy, and 4c and 4d were conducted with another BioRID II. Comparing 4a to 4b, and 4c to 4d, will show the BioRID II repeatability. Comparing 4a to 4c and 4d, and 4b to 4c and 4d, will show the BioRID II reproducibility. NIC and N_{km} values and their corresponding ORM-values are given in Table 6. Table 7 shows the ORM-values that correspond to the signals shown in Figure 22 to Figure 26.

Table 6.
NIC and N_{km} values and their corresponding ORM-values in percent for Set 4

Tests	NIC Values	NIC ORM	N_{km} Values	N_{km} ORM
4a vs. 4b	10.2 vs. 10.3	99	0.20 vs. 0.22	91
4a vs. 4c	10.2 vs. 11.6	88	0.20 vs. 0.24	83
4a vs. 4d	10.2 vs. 11.8	86	0.20 vs. 0.24	83
4b vs. 4c	10.3 vs. 11.6	89	0.22 vs. 0.24	92
4b vs. 4d	10.3 vs. 11.8	87	0.22 vs. 0.24	92
4c vs. 4d	11.6 vs. 11.8	98	0.24 vs. 0.24	100

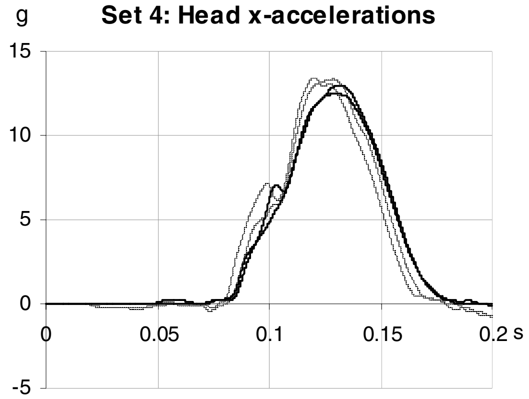

Figure 22. Head x-accelerations for all tests in Set 4. The two thick lines correspond to test 4a and 4b, and the two thin lines to test 4c and 4d.

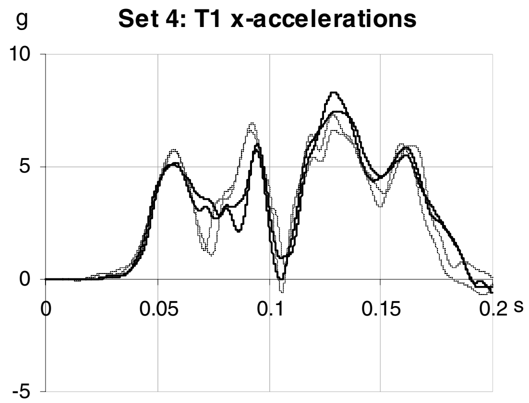

Figure 23. T1 x-accelerations for all tests in Set 4. The two thick lines correspond to test 4a and 4b, and the two thin lines to test 4c and 4d.

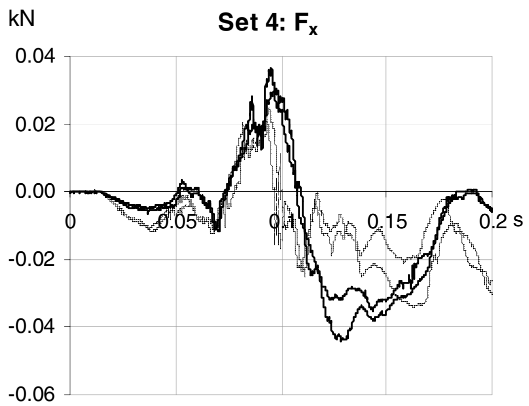

Figure 24. F_x for all tests in Set 4. The two thick lines correspond to test 4a and 4b, the two thin lines to test 4c and 4d.

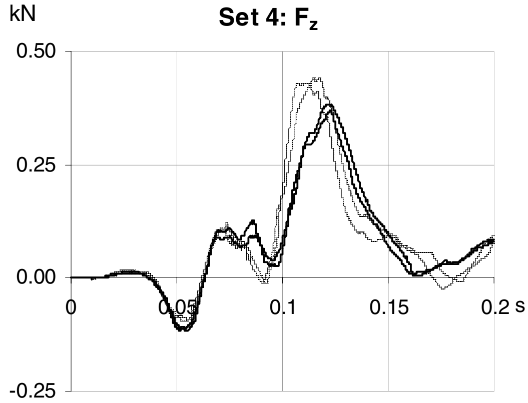

Figure 25. F_z for all tests in Set 4. The two thick lines correspond to test 4a and 4b, and the two thin lines to test 4c and 4d.

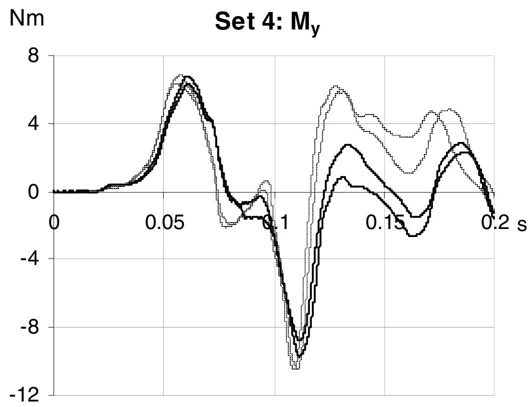

Figure 26. M_y for all tests in Set 4. The two thick lines correspond to test 4a and 4b, the two thin lines correspond to test 4c and 4d.

Table 7.
ORM-values in percent for some of the signals evaluated in Set 4

Signals		4a vs. 4b	4a vs. 4c	4a vs. 4d	4b vs. 4c	4b vs. 4d	4c vs. 4d
Head x-acc	Max peak	96	97	97	93	93	100
	Max peak time	100	98	91	98	92	93
	Curve shape	88	77	67	78	70	81
T1 x-acc	Max peak	90	93	98	84	88	95
	Max peak time	99	71	99	72	100	72
	Curve shape	86	74	72	73	71	76
Fx	Max peak	89	77	67	68	60	88
	Max peak time	99	97	95	99	97	98
	Curve shape	76	42	38	40	34	59
F_z	Max peak	96	87	89	84	86	98
	Max peak time	99	97	92	96	92	95
	Curve shape	84	66	58	62	56	74
M_y	Max peak	92	92	100	100	92	92
	Max peak time	100	95	91	95	91	96
	Min peak	90	84	84	93	93	100
	Min peak time	100	99	97	99	97	98
	Curve shape	61	46	39	35	31	66

CONLUSIONS

The Objective Rating Method (ORM) was applied to twenty-six pairs of tests in order to assess the BioRID II repeatability and reproducibility. The tests were conducted at two crash-test sites. Four BioRID II dummies, five different seats, and three crash pulses were used. Both criteria and dummy readings were compared. The BioRID II repeatability, in terms of ORM-values, ranged from 83 to 90% with a median value of 88%, and the reproducibility ranged from 74 to 78% with a median value of 77%. Based on component tests with the Hybrid III, TNO/TASS has stated that high correlation is 65% or above. Hence, the BioRID II repeatability and reproduce-ibility are very high.

REFERENCES

[1] Davidsson, J., Lövsund, P., Ono, K., Svensson, M. Y., Inami, S. (1999) *Proceedings of the 1999 International IRCOBI Conference on the Biomechanics of Impacts*, September 23-24, Sitges, Spain, pp. 165-178.

[2] Philippens, M., Cappon, H., van Ratingen, M., Wismans, J., Svensson, M., Sirey, F., Ono, K., Nishimoto, N., Matsuoka, F. (2002) Comparison of the Rear Impact Biofidelity of the BioRID II and RID2. *Stapp Car Crash Journal 46*, pp. 461-476.

[3] Roberts, A. K., Hynd, D., Dixon, P. R., Murphy, O., Magnusson, M., Pope, M. H. (2002) Kinematics of the Human Spine in Rear End Impact and the Biofidelity of Current Dummies. *Proceedings of Vehicle Safety 2002 ImechE Conference*, May 28-29, London, United Kingdom, Paper No. 2002-04-0067.

[4] Willis, C., Carroll, J., Roberts, A. (2005) An Evaluation of a Current Rear Impact Dummy against Human Response Corridors in both Pure and Oblique Impact. *Proceedings of the 19th International Technical Conference of the Enhanced Safety of Vehicles*, June, Washington, DC, USA, Paper No. 05-0061.

[5] Yaguchi, M., Ono, K., Kubota, M., Matsuoka, F. (2006) Comparison of Biofidelic Responses to Rear Impact of the Head/Neck/Torso among Human Volunteers, PMHS, and Dummies. *Proceedings of the 2006 International IRCOBI Conference on the Biomechanics of Impacts*, September 20-22, Madrid, Spain, pp. 183-197.

[6] Boström, O., Fredriksson, R., Håland, Y., Jakobsson, L., Krafft, M., Lövsund, P., Muser, M., Svensson, M. Y. (2000) Comparison of Car Seats in Low Speed Rear-End Impacts using the BioRID Dummy and the New Neck Injury Criterion (NIC). *Accident Analysis and Prevention*, Vol. 32, No. 2, pp. 321-328.

[7] Muser, M., Hell, W., Schmitt, K.-U. (2003) How Injury Criteria Correlate with the Injury Risk – A Study Analysing Different Parameters with Respect to Whiplash Injury. *Proceedings of the 18th International Technical Conference of the Enhanced Safety of Vehicles*, May 19-21, Nagoya, Japan, Paper No. 68-W.

[8] Eriksson, L., Kullgren, A. (2006) Influence of Seat Geometry and Seating Posture on NIC_{max} Long-Term AIS 1 Neck Injury Predictability. *Traffic Injury Prevention*, Vol. 7, pp 61-69.

[9] Euro NCAP – The Dynamic Assessment of Car Seats for Neck Injury Protection. Version 2.3 Final Draft. Status by 11/05/06.

[10] Hovenga, P. E., Spit, H. H., Uijldert, M., Dalenoort, A. M. (2005) Improved Prediction of Hybrid-III Injury Values using Advanced Multibody Techniques and Objective Rating. Proceedings of the SAE 2005 World Congress & Exhibition, April, Detroit, MI, USA. Paper No. 05AE-222.

[11] Davidsson, J., Flogård, A., Lövsund, P., Svensson, M. Y. (1999) BioRID P3 – Design and Performace Compared to Hybrid III and Volunteers in Rear Impacts at dv = 7 km/h. *Proceedings of the 43th Stapp Car Crash Conference*, October 25-27, San Diego, Ca, USA, Paper No. 99SC16, pp. 253-265.

[12] Ishikawa, T., Okano, N., Ishikura, K., Ono, K. (2000) An Evaluations of Prototype Seats using BioiRID-P3 and Hybrid III with TRID neck. *Proceedings of the 2000 International IRCOBI Conference on the Biomechanics of Impacts*, September, Montpellier, France, pp. 379-391.

[13] 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*, pp. 473-488.

[14] Adalian, C., Sferco, R., Fay, P. (2005) The Repeatability and Reproducibility of Proposed Test Procedures and Injury Criteria for Assessing Neck Injuries in Rear Impact. *Proceedings of the 19th International Technical Conference of the Enhanced Safety of Vehicles*, June, Washington, DC, USA, Paper No. 05-0340.

[15] Hartlieb, M. (2006) Study on Repeatability and Reproducibility of BioRID-Dummy Measurements for Whiplash Assessment. *Presented at 7th BioRID Users' Meeting*, March 1-2, Kaiserslautern, Germany, URL: http://groups.iihs.org/bioridusers/

[16] Ishii, M., Ono, K., Kubota, M. (2006) Factors Influencing the Repeatability and Reproducibility of the Dummies Used for Rear-End Impact Evaluation. *Proceedings of the 2006 International IRCOBI Conference on the Biomechanics of Impacts*, September 20-22, Madrid, Spain, pp. 405-408.