14

INJURY BIOMECHANICS RESEARCH
Proceedings of the Thirty-Fourth International Workshop

A Possible Statistical Biofidelity ATD Evaluation Scheme

G. S. Nusholtz, T. P. Hsu, and L. C. Byers

ABSTRACT

ISO/TR9790 has been in existence for some years for evaluating the biofidelity of side impact ATDs (dummies) (ISO/TR 9790, 1999). NHTSA recently generated a new method for creating biofidelity corridors. The method was different from the ISO method, by incorporating statistics and the time relationships into the evaluation equation and automating the process (Maltese et al., 2002). Although both the ISO and NHTSA methods exhibit a number of strengths, they also have weaknesses. This paper attempts to build on these two methods to develop an ATD assessment method which offers added objectivity and is based on a statistical process. Improvements were explored in several key areas to address the existing numeric issues (Hsu et al., 2005). This process bases itself on the statistical correlations between the post-mortem human subject (PMHS) data. Detailed validation of the scheme is performed using PMHSs as "pseudo dummies". A simple formula is proposed for ranking the biofidelity of the dummy, resulting in a score from 0-10, with 10 being the best.

INTRODUCTION

Since the 1950's various mechanical human surrogates, or Anthropomorphic Test Devices (ATD), have been used for assessing the potential for injury in vehicle crash tests. These ATD crash test dummies have become more sophisticated, complex, and potentially more human-like through the years, but still provide only very limited estimations of what might occur in a real life crash. In order to improve this prediction, efforts have been made through the years to make ATDs more biofidelic. However, an omni-directional dummy has not been developed. Instead, dummies have been created for each type of impact, resulting in a variety of different ATDs in frontal, side, and rear impacts. For some of these impact types, a whole family of ATD sizes has been developed.

Each of these ATDs has its own set of performance requirements, calibration procedures, and response corridors that have been developed in an attempt to make the dummy better mimic a human, as well as ensuring repeatability of responses. Many attempts have been made to determine the level of bio-fidelity. Tests have been performed using PMHSs to gather information on what injury response would be in certain

loading situations. The difficulty however, lies in how to correlate the findings from these tests with those of the ATDs, i.e., determining how accurately an ATD crash test dummy represents "real" human responses.

In 1989, the International Standards Organization (ISO) first published ISO/TR 9790, which defined a biofidelity evaluation approach for side impact dummies (Figure 1). Defining a method to standardize the determination of biofidelity was a big step forward. However, the actual method required some level of subjectivity and the resulting corridors were large, allowing the acceptability of a large amount of variation in the results.

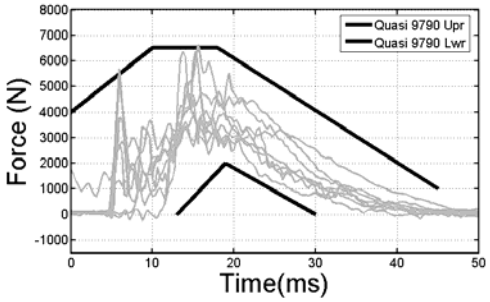

Figure 1: Typical ISO Bio-Fidelity Corridor.

In an effort to reduce the subjectivity and improve upon the ISO method, Maltese et al. (2002), at the National Highway Traffic Safety Administration (NHTSA), published a new method for creating biofidelity corridors. This method used a statistical cumulative variance approach to align the signals, which were then averaged. A corridor was then automatically created, with its boundaries defined as plus and minus one standard deviation from the mean of the aligned signals. This created a tighter corridor, which in general, better resembled the shape of the test data curves (Figure 2).

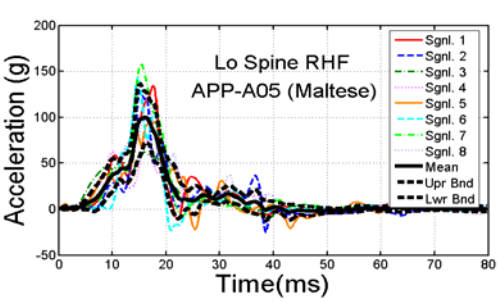

Figure 2: Typical Maltese Bio-Fidelity Corridor.

Although this method appeared to remove some of the human intervention found in the ISO method, it too had shortcomings (Hsu and Nusholtz, 2005). The selection of the standard signal with which to align the others still involved some subjectivity. The method for aligning the curves involved variability in time shifting, which could destroy relative timing information. In some situations the resulting corridor lacked physical meaning.

This paper presents a statistical, correlation-based method that builds upon the work incorporated into the ISO and Maltese methods. One notable difference in this new method is that it does not generate physical corridors. Instead it examines and compares the magnitude, shape, and phase relationships of the curves to determine the level of similarity. It then calculates a simple biofidelity score, based on these cross-correlation comparisons. This method avoids the issues caused by subjective evaluation, time shifting, and variable time history lengths. It is fully automatic and updatable. This biofidelity score is more statistics based, straightforward, and should be more representative of actual biofidelity than the existing methods.

CROSS-CORRELATION BASED APPROACH

As discussed, current dummy evaluation methods have areas that could be improved. In an earlier study (Hsu and Nusholtz, 2005), several areas for improvement were identified: including the use of a correlation method to better preserve the signal characteristics and to resolve issues resulting from time shifting, manual standard curve selection, and inconsistency in the integration time period. The proposed method aims at forming a more objective, scientific, statistically meaningful and easily applicable ATD assessment alternative.

Areas of Improvement

This approach attempts to address the following areas key to a broader scientific method in evaluating the side impact ATD:

- Incorporation of statistical correlations
- Reduction of manual intervention
- Incorporation of complete time history
- Reduction of numeric issues
- Automation of the process and improvement of process robustness

The approach proposed does not require a fixed set of PMHS data (one or greater is required). Rather, a continuously updatable set of PMHS data is used. The scheme does not shift the data as is done in Maltese's method. It is believed that the correlation method will take care of the relative timing information by using phase correlation coefficients. By eliminating time shifting, the potential destruction of relative timing information is avoided.

Steps

A flow chart of the proposed process is shown below in Figure 3. The steps to obtain the biofidelic score of an ATD consist of mass-scaling the PMHS data, scrutinizing the data using the momentum conservation theorem, calculating the magnitude, shape, and phase correlations of the PMHSs and the dummy, comparing the correlations of the dummy to the averages of that of the PMHSs (magnitude and shape), obtaining the relative phase differences between different body regions of the dummy and of each PMHS, and calculating the biofidelity score of the dummy using a multifactor based formula.

Data used in the study are from NHTSA's biomechanical research program portfolio on its public websites. Figures 4 and 5 show some typical plate force signal traces for PMHSs and dummies in sled tests. The data was mass-scaled to account for the different sized PMHSs. Since no time shifting is performed, all of the relevant signal timings are preserved.

Mass-Scaling PMHS Data

The PMHS data is mass-scaled using Eppinger's technique (Eppinger et al., 1984). The scaling process is described by Maltese et al. (2002).

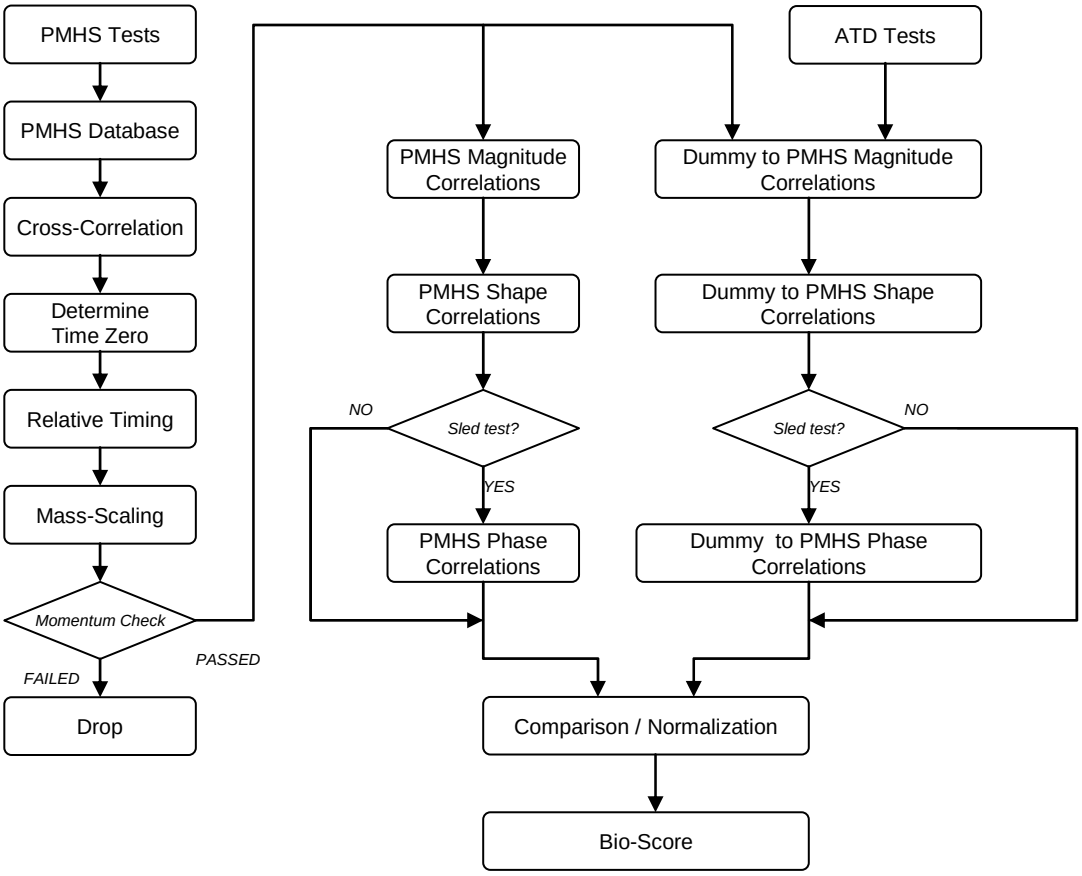

Figure 3: Flow Chart Of The Proposed ATD Evaluation Scheme.

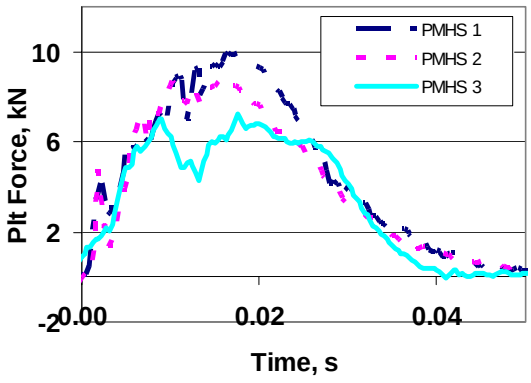

Figure 4: Typical PMHS Test Time Histories.

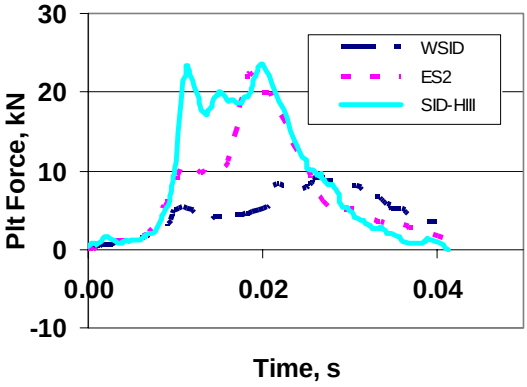

Figure 5: Typical Dummy Test Time Histories.

Data Quality Check

Before the correlations are calculated, corrupted signals need to be identified and removed to ensure the quality of the process. ISO/TR9790 incorporates a similar process by removing from the data sets the PMHSs which sustained severe rib fractures. In the approach herein, data are scrutinized using the momentum conservation theorem. The process is done to the force data, based on the theory that the summation of the force over time for a particular test condition should be relatively consistent from test to test. The same is true for the acceleration data, assuming equivalent masses can be considered constant and then applying Newton's theorem, ***F=m*a***. This way, the contamination due to instrumentation malfunction or improper calibration can be singled out easily. Since the energy inputs are the same for the group of PMHSs under the same test conditions, the integration of the response time histories from that group should yield the same value over time based on the momentum conservation theorem (Equation 1):

$$m * V = \int F dt \qquad (1)$$

i.e., for a set test condition, the velocity and the integration of the force over time should yield the same results. Those PMHSs that deviate from the majority of the group when integrated indicate a data collection issue, i.e. they either have different momentum, incorrect set-ups, an error in the data acquisition process due to miscalibration, a bad connection, or a static interference issue. The signals whose integrations deviate from the majority of the group are dropped, e.g., an arbitrary 20% has been chosen as the threshold for data elimination. Those having greater than 20% deviations from the group mean are considered to be outliers, or bad data. Only the data meeting the momentum conservation equation are used for the subsequent biofidelity evaluation. For the purpose of illustration, a set of thorax rigid plate high speed force signals is plotted in Figures 6 and 7. Five of the six tests in the graph have similar momentum, while the one with the dashed line has distinctly different integration results. All data except that test are then used for the correlation analysis.

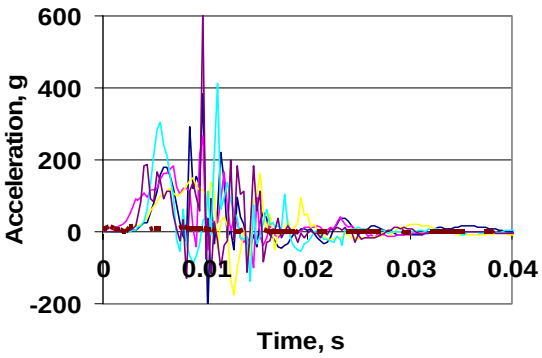

Figure 6: PMHS Signals Used For Dummy Evaluation Before Integration And Drop Of Bad Data.

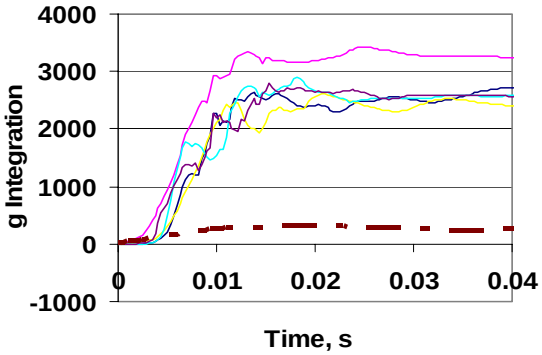

Figure 7: PMHS Signals Used For Dummy Evaluation After Integration.

PMHS Inter-Correlations

After the clean-up process, the correlation baseline from the PMHSs can be established. The cross-correlations between the PMHSs themselves are calculated using the following methods. Three quantitative indicators are utilized (Figure 8). They are magnitude, shape, and phase correlations, as described by Xu et al. (2000).

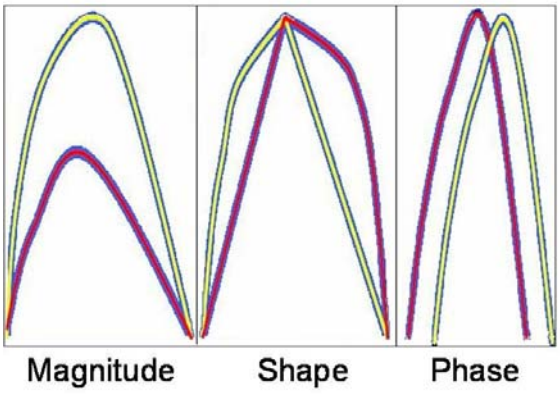

Figure 8: Schematic Representation Of Three Cross-Correlation Indicators.

Magnitude and Shape Correlations

Mathematically, the magnitude and shape correlations of the PMHSs are between 0 and 1, with one indicating that the two signals are identical (Figures 9 and 10).

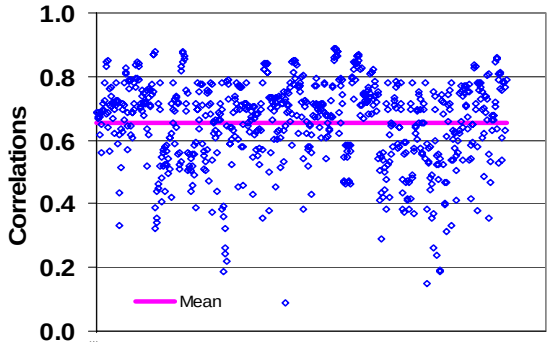

Figure 9: Typical Scatter Of Magnitude Cross-Correlation Coefficients Of PMHS Test Data.

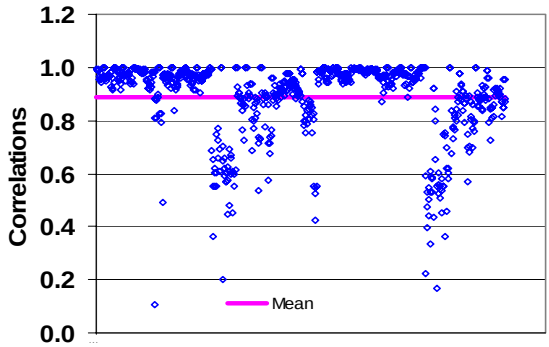

Figure 10: Typical Scatter Of Shape Cross-Correlation Coefficients Of PMHS Test Data.

The correlations are calculated in the following way to achieve reasonable and balanced numeric results. First, the one to one correlations between every two PMHSs, including itself, are calculated (Figure 11).

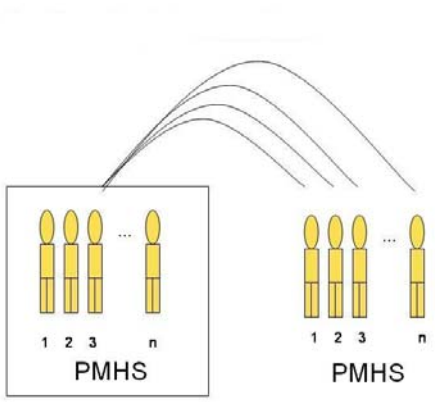

Figure 11: Magnitude And Shape Cross-Correlation Calculation.

Their sums are averaged. In averaging the sums, including or not including the auto-correlations (the cross-correlations of a PMHS with itself) yields slight differences in outcome, but is believed to be minimal (Figures 12 and 14).

Tables 1 and 2 show some correlation calculations from PMHS data in 8.9 mph tests using the PHF (Padded High speed Flat/no offset) test condition. The average of all PMHSs is shown in the top row, and the PMHS with the worst correlation for each body region (not necessarily the same PMHS for all body regions) is shown in the second row. As mentioned above, those time histories not satisfying the momentum conservation guidelines are already excluded from the calculation.

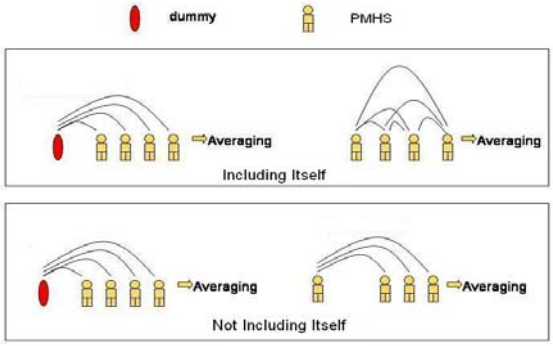

Figure 12: Two Schemes For Collective Correlation Calculations.

Table 1. Example Magnitude Correlation Results For Three Different Body Regions (Not Normalized).

BIO-FIDELITY BASED ON MAGNITUDE - PHF			
	Thorax	Abdomen	Pelvis
PMHS(ave)	0.8185	0.9164	0.8841
PMHS(wst)	0.7265	0.7104	0.8650
SID	0.2683	0.6232	0.5762
ES-2	0.5950	0.8160	0.7198
WSID	0.7212	0.7246	0.8729

Body Region Phase Correlations

While the time history magnitude and the time history shape correlations are based on a PMHS local body region, the phase relationships are compared between different anatomical regions (Figure 13). Phases between the different body regions of each PMHS are also averaged and the duration calculated. The relative timing between body regions in a crash is critical for representing human body kinematics during an impact event.

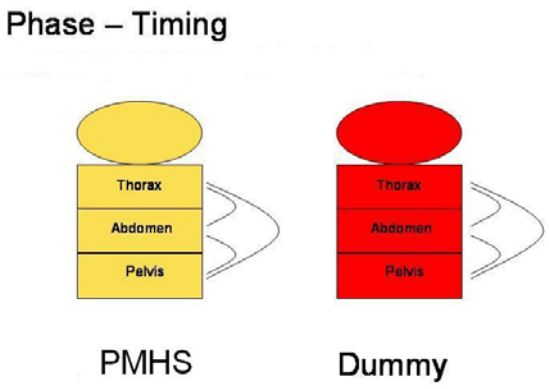

Figure 13: Schematic Representation Of Phase Correlation Calculation.

Dummy to PMHS Correlations

As the next step, a dummy's correlations to each and every PMHS are calculated similarly to the way the inter-PMHSs correlations are calculated. Some earlier/prototype test data for SID, ES-2, and WorldSID are used as an example.

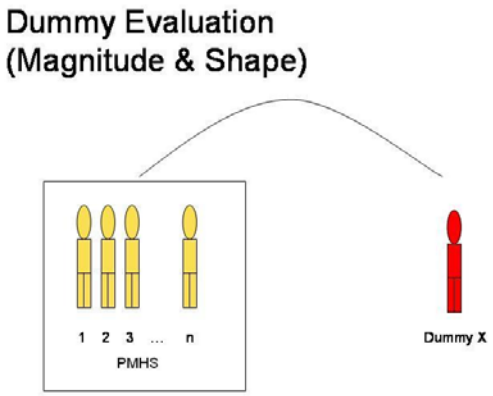

Figure 14: The Dummy Time History Is Checked Against Each PMHS Time History In The Process.

The average of the magnitude, shape and phase correlations of the dummy to each of the PMHSs is obtained. These results are shown in Tables 1, 2, and 3, in the bottom three rows.

To check the dummy's biofidelity, theoretically either the worst PMHS performer, the best PMHS performer, or the average of a PMHS group can be used as the threshold. Which is more appropriate, or more truly reflects the dummy's biofidelity, is yet to be determined. Nevertheless, the procedures are the same for either method. Only the results from the averaged PMHSs are shown in Table 1. In the following discussion, the average PMHS method is used for the purpose of describing the process. If using the best or worst PMHS is deemed to be more appropriate, it can be easily implemented without the need to change the formula.

Calculate the Bio-Score

A few variations in the dummy evaluation scheme formulation can be used, as long as the main objective remains to effectively measure the closeness of the ATD's responses to those of the PMHSs. In the proposed approach, after the correlations of inter-PMHSs and between dummy and PMHSs are calculated, summed, and averaged, the ratios of the two averages are used for the biofidelity score calculations. Equation 2 is proposed for that purpose. In Equation 2, the scaling factor of 10 is used to yield a score of 0 to 10.

The magnitude and shape correlations are normalized by dividing the average PMHS to dummy correlation by the average PMHS to PMHS correlation (Table 2). For the phase correlation contribution, the coefficients are normalized according to Equations 5, 6, and 7. They represent results as a function of total duration, as well as the time lags between different body regions.

Bio-scores are calculated for each body region:

$$Bio\ Fidelity\ \ Score\ \ =\ \ [RMAG * RSHA * RPHA] * 10 \qquad (2)$$

where,

RMAG is the ratio of average magnitude cross correlations of dummy to PMHS to that of PMHS to PMHS

$$RMAG\ =\ \sum_{i=1}^{n} \frac{CorrelationM_{dummy,i}}{n} \Bigg/ \sum_{i=1}^{n}\ \sum_{j=1}^{n} \frac{CorrelationM_{i,j}}{n,n}$$

$$(1.0,\ if\ > 1.0) \qquad (3)$$

where,

$CorrelationM_{i,j}$ is the i th PMHS to the j th PMHS magnitude correlation for a local body region.

$CorrelationM_{dummy,i}$ is the dummy to the i th PMHS magnitude correlation for a local body region.

and

RSHA is the ratio of average shape cross correlations of dummy to PMHS to that of PMHS to PMHS

$$RSHA\ =\ \sum_{i=1}^{n} \frac{CorrelationS_{dummy,i}}{n} \Bigg/ \sum_{i=1}^{n}\ \sum_{j=1}^{n} \frac{CorrelationS_{i,j}}{n,n}$$

$$(1.0,\ if\ > 1.0) \qquad (4)$$

where,

$CorrelationS_{i,j}$ is the i th PMHS to the j th PMHS shape correlation for a local body region.

$CorrelationS_{dummy,i}$ is the dummy to the i th PMHS shape correlation for a local body region.

and

RPHA is the ratio of phase lags of Dummy to PMHS to that of PMHS to PMHS

$$RPHA = RPHA_{dummy} \big/ RPHA_{PMHS}$$

$$(1.0,\ if > 1.0) \tag{5}$$

$$RPHA_{dummy} = \sum_{k=1}^{m-1} \frac{Dura - abs\left(LAG_k\right)}{Dura} \Big/ (m-1) \tag{6}$$

where,

m = Number of body regions
Dura = Duration of signal
LAG_k = Time lag between the body region being evaluated and region k.

and,

$$RPHA_{PMHS} = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m-1} \frac{Dura - abs\left(LAG_k\right)}{Dura} \Big/ (m-1) \Big/ (n*n) \tag{7}$$

where,

n = Number of PMHSs

Table 2. Example Magnitude Correlation Results For Thorax (Normalized).

MAGNITUDE - Thorax PHF		
		Normalized
PMHS(ave)	0.8185	
PMHS(wst)	0.7265	
SID	0.2683	0.3278
ES-2	0.5950	0.7269
WSID	0.7212	0.8811

All three key indicators are represented in Equation 2 and are given equal weights. The formulation of the equation yields a score of 10 when the subjects to be compared are identical and a score of 0 when they are statistically completely unrelated. The evaluation thus ties the closeness of the ATD impact response time history of a given anatomical structure to that of the impact response time history of the human surrogates using a numeric score defined as the biofidelity score. If the ATD's statistical relationships to the PMHSs are equal to or greater than those between the PMHSs, the ATD's biofidelity is considered to be excellent. If it is below the statistical relationships of the PMHSs, then it is considered to have low biofidelity.

Example

The sled data for the PHF test condition (Padded, High speed /8.9 mph, Flat/no offset) from some earlier/prototype SID-3, ES-2, and WorldSID tests are used to show the process of the biofidelity evaluation scheme being proposed (Tables 3 and A1). The PMHS data are from NHTSA's website http://www-nrd.nhtsa.dot.gov/downloads/nrd-51/2002_Stapp/. The time histories from the site already have been truncated as downloaded.

The magnitude and shape correlations of the dummy to the PMHSs are then compared to the averages of the PMHSs by dividing the correlations of the dummy to the PMHSs by that of the PMHSs. In Table A1, R5C3 (Row 5 Column 3) divided by R2C3 results in R9C3, and R6C3 divided by R2C3 results in R10C3, etc. If the quotient is greater than one, one is used instead. The same is done with the shape correlations (rows 9-11, column 4, etc.). The relative phase differences between different body regions of each PMHS are obtained (timing differences shown in ms in the example, row 2 col 5), as well as the duration of each signal (time between the first zero crossing before and after the peak time, row 2 col 6). Note that those results are yet to be updated with the latest SID-3, ES-2, and WorldSID data now available.

Table 3. Example Magnitude, Shape, And Phase Correlation Results For Thorax (Not Normalized).

		Magnitude	Shape	Phase	Duration (period)
Thorax (Not Normalized)					
	PMHS(ave)	0.8185	0.9960	0.3189	53.75
	PMHS(wst)	0.7265	0.9920		
	SID	0.2683	0.9767	5.0900	
	ES-2	0.5950	0.9895	5.5800	
	WSID	0.7212	0.9872	7.0100	

Process Verification

The sensitivity of the formula parameters is studied by using PMHS data as "dummy" data. Theoretically when PMHS data are used as "dummy" data and plugged into the formula, they should yield a good or passing bio-score because they are the very data used as the baseline to form the dummy evaluation equations.

The scheme is verified in this manner using PMHSs as "pseudo dummies." First, the NHTSA biomechanic test data on NHTSA's website http://www-nrd.nhtsa.dot.gov/database/aspx/biodb/querytesttable.aspx are downloaded. Using Eppinger's scaling technique, i.e.

$$Force_{new}(\frac{75}{M})^{\frac{2}{3}} * Force_{old} \tag{8}$$

and

$$Time_{new} = (\frac{75}{M})^{\frac{1}{3}} * Time_{old} \tag{9}$$

data are scaled according to their masses provided on the website. Table 4 lists the masses of the PMHSs and the corresponding scale factors for the times and forces. Of the eight tests under that test condition, Test 3587 is not used, due to missing force channels.

Table 4. Time And Force Scaling Factors Used Based On Eppinger's Method.

Test #	Mass	Time Factors	Force Factors
3320	**74**	1.0045	1.0090
3321	**42**	1.2132	1.4719
3323	**81**	0.9747	0.9500
3580	**56**	1.1023	1.2150
3581	**45**	1.1856	1.4057
3586	**67**	1.0383	1.0781
3589	**76**	0.9956	0.9912

The scaling baseline mass is 75 kg. Tests 3320 and 3589 have a mass of 74 and 76 kg respectively and thus have minimum impact from the mass scaling, as can be seen from the table with scale factors close to 1.0.

One issue associated with time scaling is where to position the curve after scaling. The timing is distorted after the scaling, causing the curve to be expanded or contracted. The options could be to either lock in the peak time or the time zero. In this study, the peak time is used. The positioning will have an impact on the phase factor later on in the bio-score formula. After the data scaling, signals are resampled to 3200 Hz, with a time step of .3125 millisecond.

Another issue in the process is the channel identification. Since the downloaded channels are not clearly labeled, the data have to be compared to the data downloaded from the other website, http://www-nrd.nhtsa.dot.gov/downloads/nrd-51/2002_Stapp/, to identify the thorax, abdomen and pelvis forces to be used. The associated channel numbers are listed in Table 5.

Table 5. Channel Number Line-Ups Identified.

Test #	Mass	Thorax Plate Force	Abdomen Plate Force	Pelvis Plate Force
3320	**74**	5	7	9
3321	**42**	5	7	9
3323	**81**	5	7	9
3580	**56**	5	7	9
3581	**45**	5	7	9
3586	**67**	5	7	9
3589	**76**	12	121	122

Some discrepancies were noticed after the scaling when compared to the data from the website http://www-nrd.nhtsa.dot.gov/downloads/nrd-51/2002_Stapp/. While most scaled data are consistent, Test 3321 is off by 18% (mass 42kg). Test 3580 is off by 10% (mass 56kg) and Test 3581 by 16% (mass 45kg). Those tests are found to all have a mass well under 75 kg, the scaling baseline mass used.

Data timing is kept as downloaded from the NHTSA website. The time shift by NHTSA as seen in the data from the website http://www-nrd.nhtsa.dot.gov/downloads/nrd-51/2002_Stapp/, was calculated by measuring the distances between the peaks of two corresponding curves. The amount of time shift is shown in Table A2. There seems to be minimum time shifting between body regions within each test in this particular test case, with most of them having less than 1 ms of time shift, except for Test 3323.

The time duration to be used in Equations 6 and 7 proves to be a little challenging. Some signals do not have a distinct zero crossing, primarily due to static, requiring a different method to be used for determining the time duration. An example signal is shown in Figure 15. A 5% peak force magnitude has to be used instead of the zero crossing, to define the starting time.

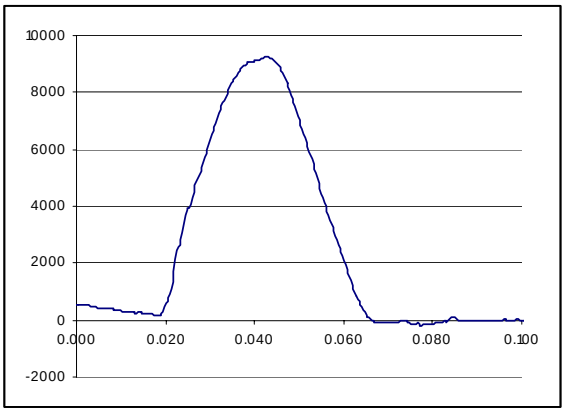

Figure 15: Starting Time Issue For A Signal Without A Zero Crossing; 5% Peak Force Is Used To Define The Starting Time Instead.

Data for Test 3581 were incomplete, perhaps due to the data length limit in NHTSA's system. Again, data from the other NHTSA website were used to extend the data (in most cases, extending the amount of data to the end), to obtain a complete time history. Other modifications include adding zeroes before or after the signal to maintain data point consistency for the entire data set.

The verification results are shown in Table A3. As expected, most of those pseudo dummies achieve high biofidelity scores.

As a reference, Maltese's corridors (the mean, upper and lower bounds) are also used as "pseudo dummies" and their bio-scores calculated (Table A4). Note that the phase factor in the bio-score formula is set to 1.0 in this case for the reason that phase lags are no longer valid due to time shift in Maltese's method. The significance of the final score has to be viewed as just an approximation.

A careful examination of the results shown in Tables 2, A1, and A3 indicates that there is a discrepancy when truncated data are used, despite the fact that identical tests are used in the calculation. The average PMHS correlations shown in Table A1 (R2C3) are inaccurate and different from Table A3 (R2C3).

DISCUSSION

Differences exist between ISO9790, Maltese's method, and the cross-correlation based approach proposed by this study. ISO9790 does not shift the signals. Maltese's method shifts the time histories based on the minimum cumulative variance relative to a master time history. ISO corridors often contain all the normalized response data within its corridors while the Maltese corridors use the signal mean plus or minus one standard deviation as the upper and lower boundaries. Three key differences between ISO9790, the Maltese method, and the approach herein are the algorithms used, the corridor definitions, and the way dummies are judged. Table A5 gives a brief summary of differences between the three evaluation schemes.

SUMMARY

A cross-correlation based evaluation scheme is proposed. The statistical characteristics of the relationship between PMHS and ATD impact response time histories are used to evaluate the ATD's biofidelity. The evaluation is done by determining if the ATD impact response time history of a given anatomical structure is statistically similar to that of the impact response time history of the anatomical structure in the PMHSs used for comparison. Three key parameters are used: magnitude correlation, shape correlation, and the phase relationship between different anatomical regions. The data relevancy is determined by kinematical factors such as conservation of momentum. The proposed approach eliminates the requirements for time shifting. The process is similar to the cumulative variance technique used in Maltese's method. It continues the work of Maltese with respect to reducing the human intervention in the existing biofidelity rating schemes.

Using this proposed biofidelity evaluation scheme, if an ATD's statistical relationships to the PMHSs are equal to or greater than the statistical relationships between PMHSs, it is considered to be biofidelic. If the statistical relationships are in the same ranges as that of the PMHSs, it is considered to be acceptable. If it is below the PMHSs, it is considered to have low biofidelity.

The process discussed in this paper is merely the framework of a side impact dummy evaluation scheme. Complete evaluation of a particular dummy requires additional work to finalize and test the scheme. Also, the proposed scheme at the time of this publication is not ready to be used for dummy design guidance. In other words, although the proposed approach certainly provides a tool for dummy evaluation, it does not provide provisions for dummy development targets. There is no corridor or curve to design a dummy to, as the ISO and Maltese methods have. Additional work will be needed to develop a similar design tool.

FURTHER WORK

More work remains to be done before this approach will be in its final form. At this time, a complete evaluation of a dummy is yet to be performed. On one hand, a more comprehensive PMHS database is needed. Without that, a reasonable statistical meaning of the scheme can not be achieved. Fortunately, thanks to the approach's flexibility, the scheme can be easily updated as additional data becomes available. On the other hand, how the correlation method should be formulated to achieve the best representation of the dummy's biofidelity remains to be further explored. How the weights should be applied to achieve the best balance of all the relevant factors in the formula (i.e. how the three correlation indicators should be weighed and combined and whether their product or summation should be used) remains to be answered. In addition, where to draw the line between the acceptable or not-acceptable ATDs is also somewhat subjective in the proposed approach. Whether a physical corridor or a score should be used as the rating tool remains to be

decided. Whether the power statistics, T-square, or some other approach should be used is to be studied as well. Whether the test data should be screened for adequacy, and how that should be done, is yet to be agreed upon by researchers in the field. All in all, there is a lot of work yet to be done, but this proposed ATD evaluation scheme provides a promising alternative in applying correlation tools in side impact dummy biofidelity evaluation.

ACKNOWLEDGEMENTS

Sincere appreciation goes to Bruce Donnelly and Kevin Moorhouse of VRTC for their support in reviewing the material and providing valuable insights, to Matt Maltese for his insights and valuable suggestions, and to Heather Rhule for providing the electronic data for the study.

REFERENCES

EPPINGER, R. H., MARCUS, J. H., and MORGAN, R. M. (1984). Development of dummy and injury index for NHTSA's thoracic side impact protection research program. Proc. SAE Government/industry Meeting, pp. 983-101 1. Society of Automotive Engineers, Warrendale, PA.

HSU, T. P. and NUSHOLTZ, G. S. (2004). 31st Annual International Workshop on Human Subjects for Biomechanical Research.

HSU, T. P and NUSHOLTZ, G. S. (2005). Considerations of Bio-fidelity Corridors for Lateral Impacts, SAE Technical Paper Series 2005-01-0308.

IRWIN, A. L., SUTTERFIELD, A., HSU, T. P., KIM, A., MERTZ, H. J., ROUHANA, S. W., and SCHERER, R. (2005). Side impact response corridors for the rigid flat-wall and offset-wall side impact tests of NHTSA using the ISO method of corridor development. Occupant Safety Research Partnership, Mich. 34 p. Stapp car crash journal. Vol. 49, p. 423-456.

ISO/TR9790 (1999). Road Vehicles Anthropomorphic Side Impact Dummy - Lateral Impact Response Requirements to Assess the Bio-fidelity of the Dummy. International Standards Organization, American National Standards Institute, NY, NY.

MALTESE, M. R., EPPINGER, R. H., RHULE, H. H., DONNELLY, B. R., PINTAR, F. A., and YOGANANDAN, N. (2002). Response corridors of human surrogates in lateral impacts. National Highway Traffic Safety Administration, Washington, D.C./ Wisconsin Medical College, Milwaukee/ Veterans Affairs Medical Center, Milwaukee, Wisc. 31 p. Stapp Car Crash Journal, Vol. 46, p. 321-351.

RHULE, H. H., MALTESE, M. R., DONNELLY, B. R., EPPINGER, R. H., BRUNNER, J. K., and BOLTE, J. H. (2002). Development of a new biofidelity ranking system for anthropomorphic test devices. National Highway Traffic Safety Administration, Washington, D.C./ Transportation Research Center, East Liberty, Ohio. 36 p. Stapp Car Crash Journal, Vol. 46, p. 477-512.

XU, L., AGARAM, V., ROUHANA, S., HULTMAN, R. W., KOSTYNIUK, G. W., MCCLEARY, J., MERTZ, H., NUSHOLTZ, G. S., and SCHERER, R. (2000). Repeatability Evaluation of the Pre-Prototype NHTSA Advanced Dummy Compared to the Hybrid III, SAE Technical Paper Series 2000-01-0165.

APPENDIX A

Table A1. Sample Results Of Bio-Fidelity Score Calculation Using Side Impact Dummies.

	1		2	3	4	5	6	7
1	PHF			Magnitude	Shape	Phase	Duration (period)	Score Scales of 0-10 10 being the most Bio-fidelic
2	Thorax		PMHS(average)	0.8185	0.9960	0.3189	53.75	
3			PMHS(worst)	0.7265	0.9920			
4								
5			SID	0.2683	0.9767	5.0900		
6			ES-2	0.5950	0.9895	5.5800		
7			WSID	0.7212	0.9872	7.0100		
8								
9		Normalized	SID	0.3278	0.9806	0.9107		2.9
10			ES-2	0.7269	0.9935	0.9015		6.5
11			WSID	0.8811	0.9912	0.8748		7.6
12	Abdomen		PMHS(average)	0.9164	0.9945	0.2706	42.19	
13			PMHS(worst)	0.7104	0.9937			
14								
15			SID	0.6232	0.9736	5.4911		
16			ES-2	0.8160	0.9830	5.4018		
17			WSID	0.7246	0.9809	6.2500		
18								
19		Normalized	SID	0.6801	0.9790	0.8755		5.8
20			ES-2	0.8904	0.9884	0.8776		7.7
21			WSID	0.7907	0.9863	0.8574		6.7
22	Pelvis		PMHS(average)	0.8841	0.9965	0.2296	40.00	
23			PMHS(worst)	0.8650	0.9934			
24								
25			SID	0.5762	0.9844	6.4732		
26			ES-2	0.7198	0.9926	4.6429		
27			WSID	0.8729	0.9850	6.6518		
28								
29		Normalized	SID	0.6517	0.9879	0.8430		5.4
30			ES-2	0.8142	1.0000	0.8890		7.2
31			WSID	0.9873	1.0000	0.8385		8.3

Due to lack of data availability, some of the results shown are not based on a complete data set.

Table A2. Time Shift (sec) per Maltese's Method.

Body Region	Test #	Peak Time 1*	Peak Time 2**	Difference	Time Shift*** wrt 3320	Time Shift wrt Thorax
Thorax	**3320**	0.02793	0.11694	0.08901	0.00000	
Thorax	**3321**	0.02608	0.14538	0.11930	0.03029	
Thorax	**3323**	0.02756	0.11805	0.09049	0.00148	
Thorax	**3580**	0.02572	0.17567	0.14995	0.06094	
Thorax	**3581**	0.02608	0.30963	0.28355	0.19454	
Thorax	**3586**	0.02682	0.27428	0.24746	0.15845	
Thorax	**3589**	0.02682	0.17161	0.14478	0.05577	
Abdomen	**3320**	0.02139	0.10971	0.08832	0.00000	0.00000
Abdomen	**3321**	0.02286	0.14209	0.11923	0.03091	0.00062
Abdomen	**3323**	0.02507	0.11596	0.09089	0.00258	0.00110
Abdomen	**3580**	0.02507	0.17558	0.15051	0.06219	0.00125
Abdomen	**3581**	0.02360	0.30842	0.28482	0.19650	0.00197
Abdomen	**3586**	0.02360	0.27088	0.24729	0.15897	0.00052
Abdomen	**3589**	0.02360	0.16748	0.14388	0.05557	-0.00021
Pelvis	**3320**	0.02219	0.11081	0.08862	0.00000	0.00000
Pelvis	**3321**	0.02219	0.14055	0.11836	0.02974	-0.00054
Pelvis	**3323**	0.01932	0.11163	0.09231	0.00369	0.00221
Pelvis	**3580**	0.02229	0.17243	0.15014	0.06152	0.00058
Pelvis	**3581**	0.02081	0.30514	0.28434	0.19572	0.00118
Pelvis	**3586**	0.02081	0.26807	0.24727	0.15865	0.00020
Pelvis	**3589**	0.02350	0.16718	0.14368	0.05506	-0.00071

* Peak Time 1 - Time (sec) at Peak from Maltese (www-nrd.nhtsa.dot.gov/downloads/ nrd-51/2002_Stapp)

** Peak Time 2 - Time (sec) at Peak NHTSA Bio Database (www-nrd.nhtsa.dot.gov/ database/aspx/biodb/querytesttable.aspx)

*** Time shift wrt Test 3320 from Time 1 to Time 2

Table A3. Sample Results Of Bio-Fidelity Score Calculation Using PMHSs As Pseudo Dummies.

	1		2	3	4	5	6
1	PHF			Magnitude	Shape	Phase	Score Scales of 0-10 10 being the most Bio-fidelic
2			PMHS (Average)	0.8453	0.9920	0.9783	
3							
4			PMHS1	0.7473	0.9936	0.9818	
5			PMHS2	0.8851	0.9903	0.9806	
6			PMHS3	0.7990	0.9906	0.9852	
7			PMHS4	0.8905	0.9928	0.9847	
8			PMHS5	0.8881	0.9929	0.9976	
9			PMHS6	0.8341	0.9940	0.9745	
10	Thorax		PMHS7	0.8731	0.9899	0.9434	
11							
12			PMHS1	0.8841	1.0000	1.0000	8.841
13			PMHS2	1.0000	0.9982	1.0000	9.982
14			PMHS3	0.9452	0.9986	1.0000	9.439
15		Normalized	PMHS4	1.0000	1.0000	1.0000	10.000
16			PMHS5	1.0000	1.0000	1.0000	10.000
17			PMHS6	0.9867	1.0000	0.9962	9.830
18			PMHS7	1.0000	0.9979	0.9643	9.623
19			PMHS (Average)	0.7844	0.9890	0.9828	
20							
21			PMHS1	0.8063	0.9902	0.9832	
22			PMHS2	0.8292	0.9897	0.9885	
23			PMHS3	0.8283	0.9847	0.9838	
24			PMHS4	0.8331	0.9888	0.9922	
25			PMHS5	0.7156	0.9909	0.9938	
26			PMHS6	0.8092	0.9918	0.9835	
27	Abdomen		PMHS7	0.6693	0.9872	0.9548	
28							
29			PMHS1	1.0000	1.0000	1.0000	10.000
30			PMHS2	1.0000	1.0000	1.0000	10.000
31			PMHS3	1.0000	0.9956	1.0000	9.956
32		Normalized	PMHS4	1.0000	0.9997	1.0000	9.997
33			PMHS5	0.9123	1.0000	1.0000	9.123
34			PMHS6	1.0000	1.0000	1.0000	10.000
35			PMHS7	0.8532	0.9981	0.9715	8.273
36			PMHS (Average)	0.8352	0.9925	0.9742	
37							
38			PMHS1	0.8531	0.9944	0.9843	
39			PMHS2	0.8699	0.9907	0.9841	
40			PMHS3	0.8187	0.9860	0.9595	
41			PMHS4	0.8251	0.9953	0.9636	
42			PMHS5	0.7850	0.9936	0.9945	
43			PMHS6	0.8691	0.9939	0.9816	
44	Pelvis		PMHS7	0.8253	0.9938	0.9518	
45							
46			PMHS1	1.0000	1.0000	1.0000	10.000
47			PMHS2	1.0000	0.9982	1.0000	9.982
48			PMHS3	0.9803	0.9934	0.9849	9.591
49		Normalized	PMHS4	0.9880	1.0000	0.9891	9.772
50			PMHS5	0.9399	1.0000	1.0000	9.399
51			PMHS6	1.0000	1.0000	1.0000	10.000
52			PMHS7	0.9881	1.0000	0.9770	9.654

Table A4. Sample Results Of Bio-Fidelity Score Calculation Using Maltese's Mean, Upper, And Lower Bounds (With Phase Factor Set To 1).

	1		2	3	4	5	6
1	PHF			Magnitude	Shape	Phase	Score Scales of 0-10 10 being the most Bio-fidelic
2			PMHS(Average)	0.8453	0.9920	0.9783	
3							
4			Mean	0.8896	0.9957		
5			Lower	0.7619	0.9927		
6	Thorax		Upper	0.7961	0.9932		
7							
8		Normalized	Mean	1.0000	1.0000	1.0000	1.000
9			Lower	0.9013	1.0000	1.0000	0.901
10			Upper	0.9418	1.0000	1.0000	0.942
11			PMHS(Average)	0.7844	0.9890	0.9783	
12							
13			Mean	0.8211	0.9944		
14			Lower	0.7216	0.9895		
15	Abdomen		Upper	0.7509	0.9911		
16							
17		Normalized	Mean	1.0000	1.0000	1.0000	1.000
18			Lower	0.9199	1.0000	1.0000	0.920
19			Upper	0.9572	1.0000	1.0000	0.957
20			PMHS(Average)	0.8352	0.9925	0.9783	
21							
22			Mean	0.8633	0.9962		
23			Lower	0.8123	0.9930		
24	Pelvis		Upper	0.8092	0.9937		
25							
26		Normalized	Mean	1.0000	1.0000	1.0000	1.000
27			Lower	0.9727	1.0000	1.0000	0.973
28			Upper	0.9689	1.0000	1.0000	0.969

Assumed values - No timing available for phase calculations.

Table A5. Comparison Of Different Side Impact Dummy Evaluation Schemes.

	ISO/TR9790	**Maltese**	**Correlation**
Evaluation	Expert evaluation	Statistical variance	Statistical correlation
Data screening	Severe rib fracture eliminated	No data exclusion	Irrelevant data eliminated through the momentum conservation theorem
Corridors	Upper and lower corridors	Mean +/- one standard deviation corridors	No physical corridors
Alignment	Manual alignment with some relative timing conservation	Alignment based on minimum variance	Alignment through correlation phase indicator
Processing	Manual processing	Automatic processing	Automatic processing
Numeric issues	No known numeric issues	The standard curve selection leading to variability; Negative corridor issue; Some irregular corridors (zero corridor width) or corridors with less physical meaning; Sometimes unstable outcome due to integration time window	No known numeric issues
Update	Update with new test data cumbersome	Updatable	Easily updatable
Manual work	More human interventions	Less human interventions	Minimum human intervention
Design guidance	Provides design guidance	Provides design guidance	Does not provide design Guidance
Ranking Scale	0 – 10 (10 best)	> 0 (the lower the better)	0 – 10 (10 best)

DISCUSSION

PAPER: **A Possible Statistical Biofidelity ATD Evaluation Scheme.**

PRESENTER: ***Guy Nusholtz, Daimler Chrysler***

QUESTION: *Erik Takhounts, NHTSA*
Guy, I couldn't help myself but noticing that some of the correlation coefficients for the dummies, as compared to the average cadaver, were greater, actually, than cadavers compared to the average cadaver. Can you explain that?

ANSWER: It's possible to have a dummy which will produce a signal which will have a higher correlation than what is in the cadavers. That is correct. Let's say I have three cadavers: two cadavers that are very close, one cadaver which is below, well below, or different than the others. So when I do the correlation, that one cadaver's going to have a very low number. Its correlation number will be fairly low. The other cadavers will be high because they're similar to each other. So when I look at an average correlation number, if I design the dummy to be close to the two that are close, or say there's four there, then it will be higher than one of the actual cadavers.

Q: And if you extend that to the total score, you have total score and the maximum's 10. You didn't show that for the cadavers, but I was wondering what is the maximum average, say, total score for your three cadavers that you use?

A: The average maximum score for the cadavers?

Q: Yes.

A: The average maximum score was above—I don't remember the exact number, but it was above the dummies.

Q: It was?

A: The cadavers were more cadaver-like than--

Q: Okay. Just wanted to make sure that they are! [laughter]

A: However, and I haven't done enough of this, but I haven't run into a situation where the cadavers were not more cadaver-like than the dummies, but I think it might be possible.

Q: It seemed that way compared to World SID so it based on your data.

A: Well if you look at the World SID does not show up as good. So World SID, if I remember, is like .7 and the cadavers are .8.

Q: Okay. Thanks.

Q: *Matt Maltese, Children's Hospital*
Guy Nusholtz, Daimler Chrysler. [laughter]

A: Well hello, Guy. How you doin'?

Q: I'm Matt Maltese from Children's Hospital. You showed correlation coefficients, average cadaver correlation coefficients for forces and they were, like, .7, .8, in that neighborhood.

A: Yeah.

Q: Did you find worse correlation coefficients for other types of signals, like accelerations, which--? Just thinking back to my experience of making corridors, it was difficult to do them or less satisfying when you did it for an acceleration because it tended to have a strange shape to it. Can you comment on that?

A: We went through all of the cadaver tests and all of the signals, accelerations, forces, whatever we could find in the NHTSA database. And if you looked at the scanner plaque, you'll see hundreds of points and those are related. Some of them get down, like I said. When they got down to .3, .2, normally it

was bad data, but you would have some correlation relationships, say with accelerations, that were .6 and .5 that mapped the conservation of energy and conservational momentum. Like I said, some of those, the shapes are so different that you can't add them. You can't just go and add the signal to try to get a mean. You'll actually distort the signal by adding them and produce a mean signal that's less than what you're individual cadaver signals are. So you lose information by adding the signals when compared to the cadaver. However, there are other signals where it's perfectly acceptable to add them. And like I said, about 10%--about 30, 40% the Maltese method, which I'm sure you're familiar with—

Q: A little, yeah.

A: Gave the best estimate in terms of maximum likelihood.

Q: Thank you.

A: You're welcome.

Q: *Richard Kent, UVA*
Maybe first just a comment: that I think we need to work a little bit on consistency of nomenclature here because my understanding of the Maltese technique—It's not really a corridor development technique. It aligns signals. And then there are various techniques like eyeball averaging or the Leslie technique which then takes those signals and makes a corridor out of them. And so, I think you hit on it. As we struggle to define what biofidelity is and what it means, and we do it in terms of corridors or correlations or something like this, but we really are not speaking a consistent language yet. And so, I think one thing, as we move forward with this and maybe would be good for an ISO group is to come up with a language that we should use for doing this because for example, Matt's study: He aligns the signals but then takes a simple, I think mean and standard deviation to define corridors. So the corridor developing from the signal alignment is not really the focus of the study. It's more aligning the signals. And so, I think we have different sorts of things going on. But anyway, that's just a comment.

A: To answer your comment: I would recommend Indo-European as the language of choice. It has been around for a while. And the second thing, I refer to Matt's method as generating a curve or it's single point, but then the corridor, the standard deviation, is used as a process of evaluation. And so specifically, his is a curve and the ISO 9790 is a corridor.

Q: Yeah. Exactly. Okay. So then maybe my question is: We're dealing with instrumentation signals that often have very complex shapes and so could you explain a little bit more how you use a single shape function to describe that shape and what is the nature of this shape function? If you've got something that looks like the top of the Grand Tetons, what's the shape factor for that and how do you compare that to something that looks like the planes in Kansas, for example?

A: Well, it's a correlation. It's similar to a correlation coefficient in a linear regression. One way to look at it—and this isn't precise, but it sort of gives you an idea: If I took one signal, one time history and somehow I aligned the two time histories to get minimum variance, and I took one and I plotted it on the y and the other I plotted on the x. If I got a straight line, then my regression coefficient would be 1 and the two signals are basically the same. But if I get deviations from that straight line, my regression coefficient goes down. So what the shape does is it says, "How close are the two in shape to each other?" And, the problem comes about when you start to have shapes, which are significantly different, then you cannot add the signals. And we've shown that if the shape is 1—I mean, you can prove this, if the shape is 1, you can add the signal and that'll give you the optimal result. If the shape is not 1, then you may not be able to add the signal. But the question is: How much off of 1 do I go before I have to use a corridor instead of using a signal? We don't know where that is and I think the approach that I'm going to be using in the near future is to try and solve it empirically. I'll just go through all the data like I did and said, "At what point does my shape coefficient degrade to a level where I cannot add the signal because it will degrade—when I add two signals—it will degrade over what it would be if I'd just taken the best cadaver signal?"

A: Okay. Thank you.

Q: *Jeff Crandle, UVA*
To follow up on that last point: Say I'm looking at a body. I'm looking at—You showed from the thorax, the abdomen or whatever. If I'm looking at something like a cross-correlation or something, can you talk a little bit about the inherent assumptions in terms of linearity of the system?

A: Well, in a way, what it's doing is it's saying how much linear it is.

Q: Yes.

A: So, that's sort of what it's telling you. In other words, if the shapes are exactly the same, do they only differ by a constant? So I can multiply some constant of one to the other to map it--

Q: But I would content with a non-linear system, they shouldn't be the same. And so your underlying measures of what's happening on a body from one region to another, with non-linearities in the system, will have changes in magnitude and phase that should not show up with a correlation.

A: They will have changes in magnitude--?

Q: From one signal to another for the same input. So in other words, what you're doing is more of an assessment of linearity than really assessment of the correlation between the two. You could have deviations purely due to non-linearity and still have the functions correlated. I mean, they could have a mapping from one to the other if they're non-linear.

A: If the two signals have different shapes and different magnitudes, it's going to be different regardless—You won't get that condition that you're stating.

Q: I believe you would for a non-linear system.

A: For a non-linear system, you're looking at the output.

Q: Sure.

A: If the two shapes are the same, there's a linear relationship between the two signals. If the two shapes are not the same, there isn't a linear relationship between the two signals.

Q: Agreed.

A: But that's what I'm saying. And then the question is: How far away from that relationship can we go and still add the signals? I don't know, but I think—

Q: I would say that's not what you're after. I would say for a given input, you could have a different body regions, to non-linear responses, you could have a deviation or an offset in the correlation coefficient and that would be okay. Let's talk about it at the break.

A: Okay.