Journal of Biomechanics 72 (2018) 258–261

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Short communication

Similitude assessment method for comparing PMHS response data from impact loading across multiple test devices

Christopher J. Dooley [a,*], Francesco V. Tenore [b], F. Scott Gayzik [c], Andrew C. Merkle [b]

[a] USAF School of Aerospace Medicine, 711th Human Performance Wing, 2510 N 5th St., Fairborn, OH 45324, United States
[b] Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Rd., Laurel, MD 20723, United States
[c] School of Medicine, Wake Forest University, Winston-Salem, NC, United States

ARTICLE INFO

Article history:
Accepted 3 March 2018

Keywords:
Corridor development
PMHS testing
Data analysis
Statistics
Correlation
Biomechanics

ABSTRACT

Biological tissue testing is inherently susceptible to the wide range of variability specimen to specimen. A primary resource for encapsulating this range of variability is the biofidelity response corridor or BRC. In the field of injury biomechanics, BRCs are often used for development and validation of both physical, such as anthropomorphic test devices, and computational models. For the purpose of generating corridors, post-mortem human surrogates were tested across a range of loading conditions relevant to under-body blast events. To sufficiently cover the wide range of input conditions, a relatively small number of tests were performed across a large spread of conditions. The high volume of required testing called for leveraging the capabilities of multiple impact test facilities, all with slight variations in test devices. A method for assessing similitude of responses between test devices was created as a metric for inclusion of a response in the resulting BRC. The goal of this method was to supply a statistically sound, objective method to assess the similitude of an individual response against a set of responses to ensure that the BRC created from the set was affected primarily by biological variability, not anomalies or differences stemming from test devices.

1. Introduction

One of the primary tools in the development of both physical models, such as anthropomorphic test devices (ATDs), and Finite Element (FE) models, in the field of injury biomechanics is the biofidelity response corridor or BRC. BRCs are used to encapsulate a range of responses for validation, as an alternative to the utilization of an individual characteristic response. Multiple techniques have been used to develop BRCs in the past, whether by bounding the set of responses (Kent et al., 2006; Nusholtz et al., 2009), finding the mean and standard deviation at each discrete time step (Gehre and Stahlschmidt, 2011), or using an objective tool such as the open source software "Correlation and Analysis" (CORA) to develop a corridor from a set of response curves. These techniques have been examined and modified methods, specific for impact and accelerative loading applications, have been proposed (Gayzik et al., 2015; Nusholtz et al., 2013; Yoganandan et al., 2004). These modifications were part of an effort to better understand human body response to under body blast (UBB) loading, characterized by high rate, short duration impact events.

An ongoing study to determine the whole body post mortem human surrogate (PMHS) response to UBB loading has utilized a proposed method specific to impact loading to develop BRCs for model development (Gayzik et al., 2015). Due to the intensive nature of whole body PMHS testing under UBB conditions, and sample size necessary to create BRCs, experiments have been conducted on both horizontal and vertical test devices, including both accelerative and decelerative designs, with the intent to combine all responses into a representative BRC. To facilitate aggregation of these data, all input conditions and positioning requirements were tightly controlled to ensure that PMHS at multiple facilities were being exposed to similar inputs. With small tested populations from a variety of test devices, an objective method for assessing similitude of responses between test rigs was used as a metric for inclusion of a response in the resulting BRC. The goal of this method is to supply a statistically sound, objective method to assess the similitude of an individual response against a set of responses to ensure that the BRC created from the set is a consistent set of responses to a similar input. The method does not suggest removal of responses, but rather reevaluation by a biomechanist with an understanding of the test device and subject.

* Corresponding author.
E-mail addresses: christopher.dooley.7.ctr@usaf.mil (C.J. Dooley), Francesco.Tenore@jhuapl.edu (F.V. Tenore), sgayzik@wakehealth.edu (F.S. Gayzik), Andrew.Merkle@jhuapl.edu (A.C. Merkle).

https://doi.org/10.1016/j.jbiomech.2018.03.010
0021-9290/

The goal of this method is to quantify the similitude of a set of responses, potentially from multiple test devices, by evaluating statistical differences in order to identify responses to reevaluate from a biomechanics or physical perspective. This method is a precursor to corridor development, to be used to screen responses before using published corridor creation techniques (Gayzik et al., 2015; Gehre and Stahlschmidt, 2011; Kent et al., 2006; Nusholtz et al., 2009; Nusholtz et al., 2013; Yoganandan et al., 2004).

2. Methods

Multiple corridor creation methods exist, a subset focusing on PMHS or ATD testing. A two-stage method has been published for comparison of multiple tests with a single ATD against themselves as well as against tests from one or more ATDs of similar design in order to assess repeatability and reproducibility of the ATD (Nusholtz et al., 2013). Another commonly used method looks uses correlation of full time series data to optimally align signals for corridor creation (Nusholtz et al., 2009). For this study the selected method utilizes a modified point-wise normalization (PWN) technique created specifically for impact and accelerative loading, focusing on the loading portion of the event (Gayzik et al., 2015).

First, given that each PMHS is employed only once, we use the term *intrarepeatability* to indicate the process of testing differences within a specific test condition (e.g. tests within an accelerative and vertical setup) using multiple (three or more) specimens, henceforth referred to as the set. A test condition is found to be *intrarepeatable* if the following criteria are met: (1) the sample size consists of three or more tests; (2) the PWN value found is greater than a threshold based on the input signal PWN, in this case 50% was selected. Fig. 1 shows example data that illustrates the degradation of signal correlation as a function of distance from the controlled input. On the left is the controlled acceleration input (the floor) with a high degree of correlation. As the energy is transmitted up the body, in this case through the calcaneus Fig. 1. Middle) and into the tibia Fig. 1. Right), the introduction of more anthropometric variables causes signal correlation to degrade. The 50% threshold was chosen for this test series heuristically due to two main features, its capability of being applied to all sensors toe to head, and its inclusivity which allowed for sufficient data for corridor creation at all locations. If a test condition has a sample size smaller than or equal to 3, this step is not executed and the data are automatically inserted into the next stage. A test condition that is identified as non-repeatable is reviewed from a biomechanics and experimental perspective to potentially identify faults in the test setup. If the test is discarded, or if intrarepeatability as a whole does not pass, then the intrarepeatability step is repeated by removing a single test to identify potential outliers in the data. This is repeated as long as the number of samples undergoing intrarepeatability analysis satisfies criteria 1.

Second, test conditions that pass intrarepeatability or that cannot undergo the intrarepeatability stage due to sample size limitations are directly fed to the *interrepeatability* and *reproducibility* stage. Here, we use a similar technique to identify potential outliers: one test at a time is removed from the set to examine how the PWN value distribution changes for the remaining responses in the set (interrepeatability). We then calculate how the removed signal correlates to the set as a whole (reproducibility). Subsequently, a two-sample non-parametric statistical test (Wilcoxon rank-sum) is used to test the null hypothesis that the samples are from the same group. Minimization of Type I errors was critical with this data set, and most biomechanical data sets, due to the small sample sizes, which helped guide the decision of using the Wilcoxon rank-sum test. Two alpha levels are explored in this study, 0.05 and 0.1. A commonly accepted statistical standard, 0.05 was the primary alpha value utilized in the analysis. To investigate sensitivity, 0.1 was used as it was expected to be a more restrictive threshold, flagging more traces as failing *reproducibility*.

3. Results

The results of this assessment are shown in Fig. 2 on a sample data set at two different significance levels, 0.05 and 0.1. The left figure shows a reduction in responses identified for reevaluation at the more inclusive significance level ($p < 0.05$) with two responses being identified, while three responses are identified at a significance level of $p < 0.1$. The traces identified at the $p < 0.05$ level both exhibit a phase lag during the initial loading phase with respect to the other traces in the set, even with all responses shifted for optimal correlation. Additionally, at the $p < 0.05$ level, one of the traces shows an evidently lower magnitude when compared to the set. One additional trace is flagged for reevaluation at the $p < 0.1$ level, this trace has the highest magnitude response of the set by ~10% but does not have the phase lag seen in the other two traces flagged. Since these are acceleration responses, they are evaluated over the range of time from 0 sec to the peak velocity.

Fig. 3 highlights the intermediate calculations performed upon each response remaining in the set after assessing intrarepeatabil-

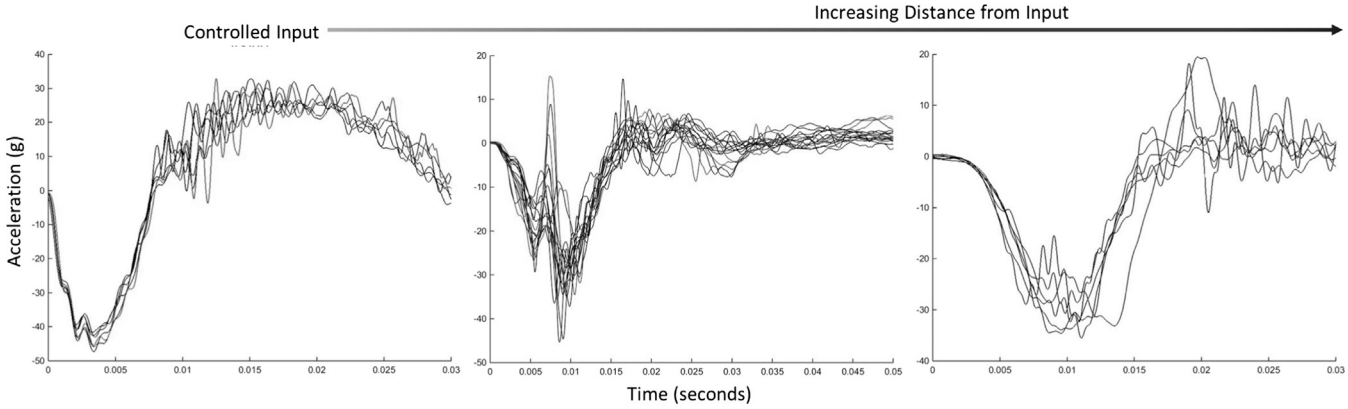

Fig. 1. An example of the degradation of signal correlation with distance from the controlled input signal. In this case, a controlled floor acceleration (left) shows a high correlation, reflected in a high PWN value. The calcaneus (middle) and tibial (right) accelerations associated with the same input are less highly correlated, illustrating the need for setting PWN thresholds for intrarepeatability based on the input correlation. The floor to lower extremity shows less degradation of correlation than the seat to pelvis/spine due to the fewer articulating joints and other dynamic variables.