Computer-Based Algorithmic Determination of Muscle Movement Onset Using M-Mode Ultrasonography

by Andrew J Tweedell, Courtney A Haynes, and Matthew S Tenan

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.
Computer-Based Algorithmic Determination of Muscle Movement Onset Using M-Mode Ultrasonography

by Andrew J Tweedell
Oak Ridge Institute of Science and Education, Oak Ridge, Tennessee

Courtney A Haynes and Matthew S Tenan
Human Research and Engineering Directorate, ARL


Approved for public release; distribution is unlimited.
## 14. ABSTRACT

The study purpose was to evaluate the use of computer-automated algorithms as a replacement for subjective, visual determination of muscle contraction onset using M-mode ultrasonography. Biceps and quadriceps contraction images were analyzed visually and with three different classes of algorithms: pixel standard deviation (SD), high-pass filter and Teager Kaiser energy operator transformation. Algorithmic parameters and muscle onset threshold criteria were systematically varied within each class of algorithm. Linear relationships and agreements between computed and visual muscle onset were calculated. The top algorithms were high-pass filtered with a 30 Hz cutoff frequency and 20 SD above baseline, Teager Kaiser energy operator transformation with a 1200 absolute SD above baseline and SD at 10% pixel deviation with intra-class correlation coefficients (mean difference) of 0.74 (37.7 ms), 0.80 (61.8 ms) and 0.72 (109.8 ms), respectively. The results suggest that computer automated determination using high-pass filtering is a potential objective alternative to visual determination in human movement science.

## 15. SUBJECT TERMS

ultrasound, algorithms, muscle movement onset, agreement, muscle contraction, visual determination

<table>
<thead>
<tr>
<th>16. SECURITY CLASSIFICATION OF:</th>
<th>17. LIMITATION OF ABSTRACT</th>
<th>18. NUMBER OF PAGES</th>
<th>19a. NAME OF RESPONSIBLE PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. REPORT</td>
<td>Unclassified</td>
<td>UU</td>
<td>Andrew J Tweedell</td>
</tr>
<tr>
<td>b. ABSTRACT</td>
<td>Unclassified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. THIS PAGE</td>
<td>Unclassified</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>19b. TELEPHONE NUMBER (Include area code)</td>
<td>410-278-5813</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
COMPUTER-BASED ALGORITHMIC DETERMINATION OF MUSCLE MOVEMENT ONSET USING M-MODE ULTRASONOGRAPHY

ANDREW J. TWEEDELL, COURTNEY A. HAYNES, and MATTHEW S. TENAN
U.S. Army Research Laboratory, Human Research & Engineering Directorate, Aberdeen Proving Ground, Maryland, USA

Abstract—The study purpose was to evaluate the use of computer-automated algorithms as a replacement for subjective, visual determination of muscle contraction onset using M-mode ultrasonography. Biceps and quadriceps contraction images were analyzed visually and with three different classes of algorithms: pixel standard deviation (SD), high-pass filter and Teager Kaiser energy operator transformation. Algorithmic parameters and muscle onset threshold criteria were systematically varied within each class of algorithm. Linear relationships and agreements between computed and visual muscle onset were calculated. The top algorithms were high-pass filtered with a 30 Hz cutoff frequency and 20 SD above baseline, Teager Kaiser energy operator transformation with a 1200 absolute SD above baseline and SD at 10% pixel deviation with intra-class correlation coefficients (mean difference) of 0.74 (37.7 ms), 0.80 (61.8 ms) and 0.72 (109.8 ms), respectively. The results suggest that computer automated determination using high-pass filtering is a potential objective alternative to visual determination in human movement science. (E-mail: andrew.tweedell.ctr@mail.mil) Published by Elsevier Inc. on behalf of World Federation for Ultrasound in Medicine & Biology.

Key Words: Muscle contraction, Visual determination, Agreement.

INTRODUCTION

In clinical biomechanics and human movement research, the timing of muscle contraction, or muscle onset (MO), is often used to examine the temporal relationship between external and internal events (e.g., electromechanical delay [Begovic et al. 2014], agonist/antagonist co-activation [Hortobagyi et al. 2009] or during the gait cycle [Powers et al. 1996]). Traditionally, surface and intramuscular electromyography (EMG) is used to measure MO; however, intramuscular EMG is an invasive procedure and surface EMG is a diffuse measure highly influenced by the activity of superficial muscle fibers (Farina et al. 2004).

The limitations of EMG have led many researchers to use musculoskeletal ultrasonography (US) to observe the mechanical displacement of deep muscles and associate this with muscle contraction (Begovic et al. 2014; Hodges et al. 2003; Pulkovski et al. 2008; Vasseljen et al. 2009). In US, transmitted and reflected ultrasonic waves (f ≥20 kHz) create a backscattering of waves resulting in a 2-D array of gray-scale values producing a typical sonogram image. In M-mode US, this image is updated hundreds of times per second to produce a time series of tissue displacement underneath the US probe. Processing of M-mode images for MO determination is often accomplished by leveraging techniques commonly applied to the EMG signal, such as high-pass (HP) filtering to remove artifact (Vasseljen et al. 2006, 2009).

In most cases, the gold-standard method for MO determination is visual analysis, where investigators determine the time point of MO (Vasseljen et al. 2006, 2009; Westad et al. 2010). Although each study has reported good reliability for visual measurements (Pulkovski et al. 2008; Vasseljen et al. 2006, 2009), this method is time consuming, is highly dependent on the experience of investigators and may not be generalizable.

Alternatively, M-mode images can be processed via computer algorithms to detect when there is a change in pixel values (indicating movement initiation) across time. For example, Dieterich et al. (2015) developed a process using a method previously applied to EMG (Li et al. 2007; Solnik et al. 2010) by pre-conditioning US gray-scale values with the Teager-Kaiser energy operator (TKEO) (Kaiser 1990, 1993) and setting specified standard deviation threshold criteria for MO. The
subsequent US-derived MO times were highly correlated (intra-class correlation coefficients [ICCs] = 0.73–0.93) with simultaneously recorded EMG-derived MO of the gluteus medius and gluteus minimus. Such automated detection techniques have the potential to replace tedious visual analyses. However, studies attempting to develop US-derived MO, whether visual or automated, often compare it with EMG-derived MO (Dieterich et al. 2015; Mannion et al. 2008; Pulkovski et al. 2008; Vasseljen et al. 2006, 2009). Given that EMG and ultrasound are two physiologically distinct measures, assuming that EMG can act as external validation of ultrasound are two physiologically distinct measures, MO may not be appropriate (Vasseljen et al. 2006). To date, there has been no attempt to validate automated US MO techniques against the gold standard for US MO determination, visual analysis.

A reliable method to effectively measure MO using US, independent of investigator bias, could potentially be useful to clinicians and researchers investigating musculoskeletal function. Computer-determined MO, as a method, provides a speed and reliability advantage over visual analysis. The primary aim of this study was to investigate computer-based MO algorithms as a viable alternative to visual determination using M-mode US for both simple and complex muscles. The secondary aim was to characterize and compare the reliability of visually determined MO using raw M-mode images or a time-series waveform of pixel data.

METHODS

Thirteen men and five women (aged 22–54 y) with no neurologic disorders or musculoskeletal injuries were recruited for this study and subsequently signed an informed consent document. The study protocol was approved by the U.S. Army Research Laboratory institutional review board, and all participants gave their informed consent in accordance with the Helsinki Declaration.

Three separate contractions for both knee extension and elbow flexion with 2.3 kg of resistance were performed in a seated upright position. Tasks were performed at a self-selected moderate speed on the verbal command of an investigator. Ultrasonic data were collected with a Sonosite Edge Ultrasound System (FUJI-FILM SonoSite, Bothell, WA, USA) using a multifrequency linear array transducer (HFL38x, 13–6 MHz, 6 cm, SonoSite). The same trained investigator (A.J.T.) performed all US assessments.

The ultrasound transducer was placed on the muscle belly parallel to the underlying muscle fibers of the vastus lateralis and biceps brachii for knee extension and elbow flexion, respectively. Image depth and overall gain were optimized for each participant (2–6 cm) to account for variation in subcutaneous fat thickness between participants and to ensure muscle tissue was captured for image analysis. After visualization, the US system was set to M-mode at a temporal resolution of 200 Hz. This is equivalent to a single US beam recording the image slice of the underlying fat, muscle and tendon architecture once every 5 ms to produce a time series image (Fig. 1a). The 8-bit gray-scale pixel data (0–255 arbitrary units [a.u.]) were encoded as a matrix of 512 columns by 250 rows of pixels and exported by USB as a DICOM JPEG to the image processing program ImageJ (Version 1.47, National Institutes of Health, Bethesda, MD, USA) and a customized R language software program (R Development Core Team 2015) using RStudio (RStudio, Boston, MA, USA) and associated packages (Gordon and Lumley 2016; Kienzle et al. 2014; Lehnert 2014; Revelle 2016; Vanderkam and Allaire 2015; Whitcker et al. 2011) for processing. Each column of the matrix represents a single point measure (at 5-ms intervals), and each row represents a time series of the architecture at that specific depth.

Visual determination

Visual determination of muscle onset was accomplished with two separate methods. Each method employed a blinded analysis by the same two independent investigators (C.A.H. and M.S.T.). Note that the trained US operator (A.J.T.) was not among the investigators performing the blinded analysis. For the first method, the investigators analyzed the raw M-mode sonogram (Fig. 1a) in ImageJ to identify the initial instance in which there was a clear indication of high-energy motion, defined as a general perturbation in pixel coloration. This method had not previously been attempted. For the second method, the pixel matrix was imported into a custom R computer program and transformed with an HP filter, as previously used by Vasseljen et al. (2006, 2009). First, each row of the pixel matrix was transformed using a zero-lag fourth-order HP (forward–backward second-order Butterworth) IIR filter with a 30-Hz cutoff frequency. The root mean square (RMS) value of all transformed gray-scale pixel values within each column was then calculated and plotted as a graphical time series (Fig. 1b). The investigators were asked to identify the first instance in the graph in which there is a clear deviation in RMS from baseline.

Computerized algorithms

Computed MO was determined by three separate classes of algorithms using RStudio: (i) a novel standard deviation (SD) technique, (ii) an HP filtering technique and (iii) a TKEO pre-conditioning technique. An iterative analysis of each class was run with systematically varying parameters for MO determination. The latency (in ms) between the determined onset and the beginning of the
M-mode trace was subsequently used to compare the time to onset between algorithms.

**Standard deviation.** The SD technique included calculation of the $z$-scores of pixels within each row, as indicators of variation in the gray-scale values over time. Then the percentage of pixels in each column (i.e., time point) in which the $z$-score was $\pm 1.96$ was calculated. MO was determined as the point when the percentage was greater than 5%, 10%, and 15% (referred to as SD 5%, SD 10% and SD 15%, respectively). This technique allows for the relative quantification and comparison of the standardized gray-scale values that lie outside the 95th percentile of a normal distribution for each time point.

**High-pass filtering.** The process for HP filtering and RMS was described under Visual Determination; for the algorithmic method, MO was chosen computationally instead of by investigators. Filtering parameters and threshold criteria were varied between 5-, 15- and 30-Hz cutoff frequencies with 5, 10 and 20 SD above baseline (referred to as “HP” with respective cutoff frequency and criteria, e.g., HP 15 Hz 10 SD).

**TKEO pre-conditioning.** For TKEO pre-conditioning (Li et al. 2007; Solnik et al. 2010), each row of the pixel matrix was transformed using the TKEO formula (Kaiser 1990; Kaiser 1993). The SD of each column then was calculated. The threshold criteria for MO were determined by previous studies as having an absolute SD of 800 or 1200 a.u. (Dieterich et al. 2015). (referred to as TKEO 800 and TKEO 1200, respectively). The SD of each column indicates variation in energy levels between time points.

**Statistics**

The systematic variability of visual determination between investigators was evaluated using the intra-class correlation coefficient ($\text{ICC}_{3,1}$) and standard error of measurement (SEM). The $\text{ICC}_{3,1}$ model measures the systematic error between only investigators of interest so the results can be used for comparison between other
methods based on a single measurement taken by each investigator (Shrout and Fleiss 1979). The visual determination method that was considered more reliable (i.e., higher ICC3,1 value) was chosen as the “gold standard.” The mean MO value of the two investigators’ measurements was used for further statistical analysis.

Given the large number of MO algorithms, poorly performing algorithms were systematically eliminated from further evaluation. First, algorithms that detected muscle onset in less than 95% of trials were removed. Second, a hierarchical cluster analysis was performed on the remaining algorithms. This analysis attempts a stepwise agglomeration of similar onset times into unknown clusters by their proximity to each other based on the Euclidean distance between them. This produces a dendrogram depicting the relative similarity between algorithms, with connected branches of the graph representing commonality. Cluster analyses are widely employed in fields such as genomics for microarray analyses to uncover common traits in very large data sets (Kaufman and Rousseeuw 2005). The three algorithms with the closest proximity (i.e., highest similarity) to the gold-standard visual determination on the dendrogram were selected for further evaluation. Third, the level of agreement between the chosen algorithms and visual determination was assessed with mean differences (bias) and precision of the estimated mean differences (95% confidence interval [CI]) using the Bland–Altman method (Bland and Altman 1986). Any missing values (i.e., no onset was detected) were treated with a pairwise deletion. The ICC3,1 values were then calculated using the R statistical package “psych” in RStudio (Revelle 2016). Additionally, the ranges of differences observed are reported for each of the top three algorithms.

RESULTS

Fifty-four and 53 trials were performed for vastus lateralis and biceps brachii, respectively. Visual determination using the filtering method produced greater ICC3,1 values (0.99, SEM = 15.3 ms) between investigators than use of the raw image (ICC3,1 = 0.91, SEM = 37.0 ms). Therefore, visual MO times from HP filtering were used for further analysis. Similarity between this visual determination and all computer algorithms with ≥95% detection rates are displayed in the dendrogram in Figure 2.

Based on their relative commonality and proximity to “visual” MO according to the cluster analysis (Fig. 2), the top three algorithms chosen for further evaluation were HP 30 Hz 20 SD, TKEO 1200 and SD 10%, with mean differences (ms) and 95% CIs of 37.7 (11.7–63.7), 61.8 (39.6–83.9) and 109.8 (85.3–134.4), respectively. Bland–Altman plots were constructed to visually assess the agreement between visual determination and each algorithm (Fig. 3). A forest plot of the mean differences and 95% CIs is provided in Figure 4 along with ICC3,1 values and ranges of differences observed.

DISCUSSION

In an effort to replace EMG with US as the main method of MO determination, previous studies often compared onset timing between the two. This is misleading because they quantify different physiologic phenomena. Theoretically, electrical activation should always precede muscle fiber movement, but many studies report trials in which US-derived MO occurred before MO derived from EMG (Mannion et al. 2008; Pulkovski et al. 2008; Vasseljen et al. 2006, 2009). Vasseljen et al. (2006) hypothesize that electrical activation of a muscle measured by EMG does not necessarily equate to muscle movement because nearby, earlier-activated muscles may cause passive motion to propagate through the US field of vision. This makes validation of any automated US MO technique against EMG MO potentially problematic, especially when attempting to distinguish between electrical and mechanical events (i.e., electromechanical delay). An appropriate comparison between methods using the same instrumentation (US vs. US) is required to determine the accuracy of any automated techniques. This is the first study to compare automated US methods against visual analyses for determining the point of muscle contraction onset in vivo. It is also a step toward a standardized method of US MO determination.

The results indicate that reasonable inter-method reliability (ICC3,1 values = 0.72–0.80) can be achieved with algorithmically computed onset using M-mode US. Taking both muscles into account, the HP filtering algorithm with a 30 Hz cutoff frequency and 20 SD threshold produced the lowest bias (37.7 ms) with a
precision (11.7–63.7 ms, 95% CI) comparable to those of the others, making it the most viable option. Additionally, visual inspection of the Bland–Altman plots (Fig. 3) reveals that the distribution of data points is similar between muscles for each algorithm, possibly indicating that MO measurements are not affected by differences in muscles. According to the agreement analysis, there is a statistical bias for algorithmic detection to occur after visual detection, as evidenced by the positive mean differences seen in Figures 3 and 4. This positive bias may be due to the algorithms themselves or to the specific threshold criteria set a priori for each algorithm. Each algorithm’s parameters were chosen based on previous methods using US or EMG. Further adjustment of these parameters may yield better or worse results, but lowering the thresholds will likely increase the chances of premature onset detection. Additionally, although the good to excellent ICC values (Fig. 4) indicate there is consistency between visual and algorithmic MO, each algorithm experienced a large range of differences from visual. For instance, the SD 10% algorithm reports a maximal error (maximum difference observed between algorithm and visual) of almost a second, whereas HP 30 Hz 20 SD and TKEO 1200 report 737.5 and 592.5 ms, respectively. These maximal errors may represent programmatic “edge cases” where the algorithm returns results that are not representative of typical performance and are largely unavoidable when implementing algorithms.

For measures that require high-temporal-resolution US to distinguish specific events, such as electromechanical delay (Begovic et al. 2014; Nordez et al. 2009), the bias observed in this study may not be acceptable for computer algorithms to replace visual determination for MO as of yet. As a limitation, the temporal resolution for M-mode in the present study consisted of 5 ms (200 Hz), whereas others using US for MO report <2 ms (<500 Hz) (Begovic et al. 2014; Nordez et al. 2009; Vasseljen et al. 2006, 2009). It is expected that with higher refresh rates, the algorithms may more precisely predict onset. However, temporal resolution appears to have little influence on manual reproducibility because both visual methods for MO had excellent inter-rater reliability, with the HP filtering method producing ICC values similar to those of previous studies (Vasseljen et al. 2006, 2009). It should be reiterated that even with excellent reliabilities and narrow SEMs between investigators (15.3 and 37.0 ms), visual analysis is a subjective measure with possible human error. Because the purpose of the present study was to develop an objective replacement for visual analysis, it is assumed any biases
Algorithmic determination of muscle movement

A. J. Tweedell et al.

reported were due to the algorithms and were not errors in visual determination. An example M-mode trace with MO time points from each algorithm and observer highlighted is presented in Figure 1.

Clinically, the absolute reliability and repeatability for algorithmic determination methods make them ideal for studies that implement repeated measures. These advantages offer potential for many experimental paradigms, from tracking performance (e.g., sports training or rehabilitation) to evaluating man-machine control systems (e.g., prosthetics). For a clinical study example, musculoskeletal function can be assessed pre- and post-orthopedic surgery using these automated US techniques. In addition, advanced US systems are now equipped with Doppler imaging which can quantify tissue velocities (Pulkovski et al. 2008; Westad et al. 2010). This type of ultrasound technology may require new or different methods and algorithms to determine MO.

CONCLUSIONS

At this time, visual determination using HP filtering remains the most effective method for overall MO determination; however, this study provides evidence that algorithmic detection using HP filtering is a viable alternative that offers researchers a more objective measure for muscle movement onset.

Acknowledgments—The authors thank all patients who participated in this study.

REFERENCES


Lehnert B, BlandAltmanLeh: Plots (slightly extended) Bland–Altman plots, Version 0.3.1. 2014. Available at: https://cran.r-project.org/web/packages/BlandAltmanLeh/BlandAltmanLeh.pdf.


