Physical Medicine and Rehabilitation Research

Research Article ISSN: 2398-3353

Quantification of elbow muscle tone from an instrumented manual stretch-reflex test

Chris A. McGibbon[1,2*], Andrew Sexton[1], Melony Jones[3] and Colleen O'Connell[2,3]

[1]Institute of Biomedical Engineering at the University of New Brunswick, Fredericton, New Brunswick, Canada

[2]Faculty of Kinesiology, University of New Brunswick, Fredericton, New Brunswick, Canada

[3]Stan Cassidy Centre for Rehabilitation, Fredericton, New Brunswick, Canada

Abstract

Objective: To evaluate a method for objectively quantifying elbow muscle tone in a clinical setting using an instrumented manual stretch-reflex test.

Methods: Seventy-nine participants with upper motor neuron syndrome (stroke, spinal cord injury, cerebral palsy and multiple sclerosis) were evaluated for elbow flexor and extensor tone using a wearable sensor system. Modified Ashworth Scale (MAS) scores of elbow flexors and extensors, and spasticity metrics derived from a uniform-jerk model during manual stretch-reflex test, were used in a linear discriminant analysis (LDA) to generate a probability based 0-10 score (.1 increment) that maps onto the MAS continuum.

Results: Sensor derived metrics correlated significantly with EMG (onset time: r^2=.7, p<.001; duration: r^2=.9, p<.001) and explained as much as 50% of the variance in therapist-rated MAS score. The LDA resulted in 73% classification accuracy, although the "gold standard" MAS rating was a considerable source of error.

Conclusions: The study demonstrates that a simple wearable sensor system in combination with a routine manual stretch-reflex test can be used to objectively quantify elbow flexor and extensor tone. These findings offer new hope of achieving objective measurement of muscle tone in the clinic.

Introduction

Upper motor neuron syndrome (UMNS) is a life-long chronic condition affecting people with brain and spinal cord injury or disease [1]. A hallmark of UMNS is muscle spasticity which interferes with residual function and may lead to painful muscle contractures if not treated appropriately [2]. Spasticity management often includes pharmacologic treatment (injections, orals, pumps, etc.) which are costly [3], and have questionable efficacy [4]. Regardless of the approach, management of UMNS relies on muscle tone assessment for correct treatment prescription [5], which should be objective, repeatable and precise.

Spasticity assessment in the clinic is currently limited to subjective rating scales such as the Modified Ashworth Scale (MAS) [6]. The MAS is based on a manual stretch-reflex test (SRT) where the patient is positioned supine, arm is relaxed, and the clinician moves the limb from extension to flexion (or vice versa) over a "one second" count. Based on the definition of spasticity as a velocity-dependent hyper-excitability of the stretch-reflex [7], spastic muscle will involuntarily contract during the passive stretch. To apply the MAS, the degree of muscle catch and prolonged resistance are subjectively rated by the clinician and reported on a 6-point scale, as shown in Table 1.

Although the SRT is easy to reproduce, the MAS rating lacks sensitivity to change and reliability is moderate to low in some studies [8,9]. MAS scores have been compared to objective assessment using mechanical devices that induce SRT motion and measure torque response (isokinetic dynamometers or other custom motorized machines) but results vary from showing very little association [10-12] to relatively good association [13-15] with stiffness measures. In studies employing EMG the reflexive component correlates with MAS score [16,17] or its variants [18].

Wearable kinematic sensors in combination with EMG have been studied during a manual stretch-reflex test (SRT) and have been shown to be feasible for use in a clinical environment [19-21] but results of these studies are difficult to compare due to different experimental approaches. Calota *et al.* [19] use multiple SRTs at varying velocity to derive the tonic stretch-reflex threshold (TSRT) [22] and found no relationship with MAS scores. Our preliminary study [20] found a positive relationship between clinical MAS rating and the magnitude of "trajectory departure" (via jerk response) during instrumented elbow SRT, that also corresponded with EMG responses, suggesting that some aspect of spasticity may be quantifiable from a single SRT at a typical testing velocity.

The present paper explores a simple framework for interpreting and analyzing kinematic signals from the instrumented SRT when the therapist is the mechanism that delivers the impulsive dynamic motion, and demonstrates how such a framework could be implemented as a potential solution to the problem of objective spasticity measurement

Correspondence to: Chris A. McGibbon, PhD., Institute of Biomedical Engineering at the University of New Brunswick, Fredericton, New Brunswick, Canada, Tel: 506-458-7098; E-mail: cmcgibb@unb.ca

Key words: *motion sensors, spasticity, muscle tone assessment, elbow muscle, upper motor neuron syndrome, modified ashworth test*

Received: November 20, 2016; **Accepted:** December 15, 2016; **Published:** December 20, 2016

Table 1. Modified Ashworth Scale rating criteria (Bohannon and Smith, 1987).

Score	Presentation
0	No increase in muscle tone
1	Slight increase in muscle tone, manifested by a catch and release, or by minimal resistance at the end of the range of motion when the affected part(s) is moved in flexion or extension
1+	Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the range of movement (ROM)
2	More marked increase in muscle tone through most of ROM, but affected part(s) easily moved
3	Considerable increase in muscle tone, passive movement difficult
4	Affected part(s) rigid in flexion and extension

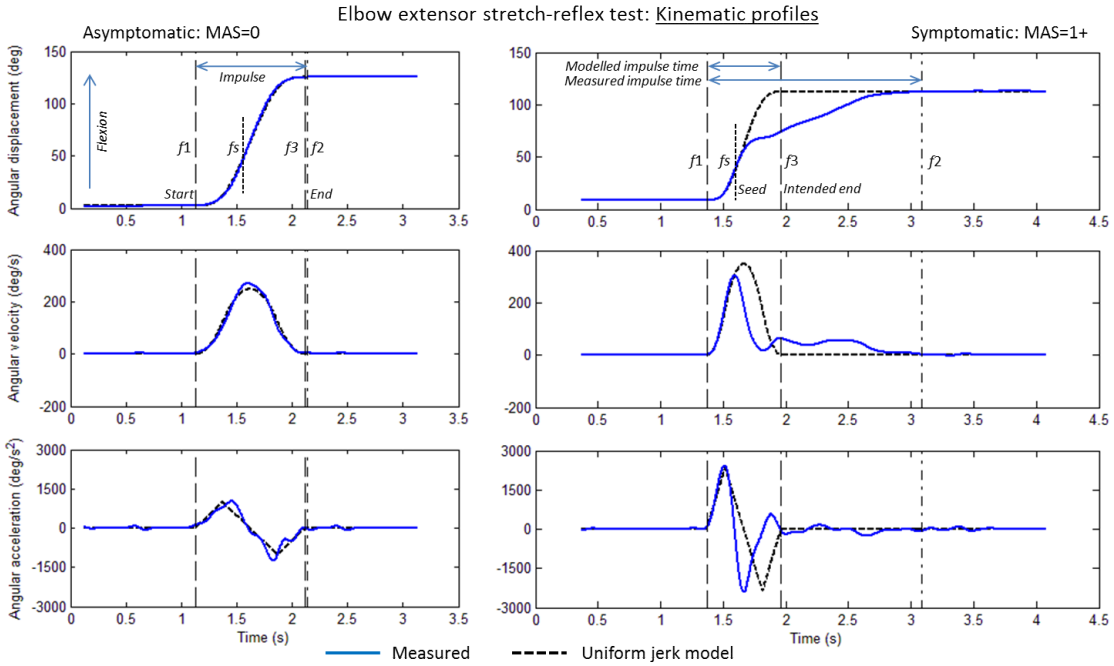

Figure 1. Elbow kinematic profiles measured during a stretch-reflex test of elbow extensors (solid line) and uniform jerk model (dashed line) for angular displacement (top), angular velocity (middle), and angular acceleration (bottom). Left panels show trajectories for asymptomatic extensor with MAS=0, and right panels show trajectories for symptomatic extensor with MAS=1+. Start time (*f*1) and end time (*f*2) of the motion, and a "seed" time (*f*s) of the impulse, are used to compute the constant-jerk curve and its integrals, and the modeled impulse end time (*f*3).

in the clinic or home. Although the approach could be applied to any clinical rating based on the SRT, we focus in this paper on the MAS due to its wide use in practice and clinical trials [23] and a general need to better understand what it measures [13].

Methods

Model framework

As introduced previously [20], the trajectory of the passively moved limb segment (forearm in this case) of the patient by the therapist during a manual SRT follows a uniform jerk profile. Jerk is the fourth kinematic or third derivative of position, and therefore represents the rate of change of acceleration. It is well established in the literature (see seminal work by Flash and Hogan [24]) that in healthy individuals voluntary movement of body segments naturally attempt to minimize jerk in order to produce a smooth velocity profile of the end-effector (hand or foot) [25]. It was therefore postulated that in order to rapidly extend the forearm through its range of motion the therapist would naturally select a uniform jerk profile to produce a rapid but smooth trajectory of the patient's limb.

A uniform jerk displacement curve fits almost perfectly the measured elbow trajectory in arms with normal tone, as shown in Figure 1 (left). However, the uniform jerk profile during passive movement of the limb will only hold true if the movement is not obstructed or resisted by the passively stretched muscles; i.e., there is no sudden or progressive force generated in response to the motion trajectory. Of course, muscle spasticity does exactly this – the stretching muscle will involuntarily contract during the trajectory causing a sudden opposing force (catch), which may release but and in some individuals the muscle activity persists throughout most if not all of the remaining movement range (prolonged resistance).

Figure 1 illustrates SRT data from an asymptomatic elbow (left panel, MAS=0) and symptomatic elbow (right panel, MAS=1+). The model input is the measured angular trajectory of the patient's joint, which is used to generate a uniform jerk profile of the therapists intended trajectory of the joint. The model's jerk profile represents two sequential uniformly positive and negative steps (not shown), thus forming the linear acceleration curve, quadratic velocity (bell shaped) curve, and cubic displacement curve, shown in Figure 1 by the dashed lines. The working hypothesis is that spasticity explains, at least in part, the difference between measured and modelled curves during the SRT.

Figure 2 shows the kinematic metrics that describe the observed departure from the model curve, which include the peak departure

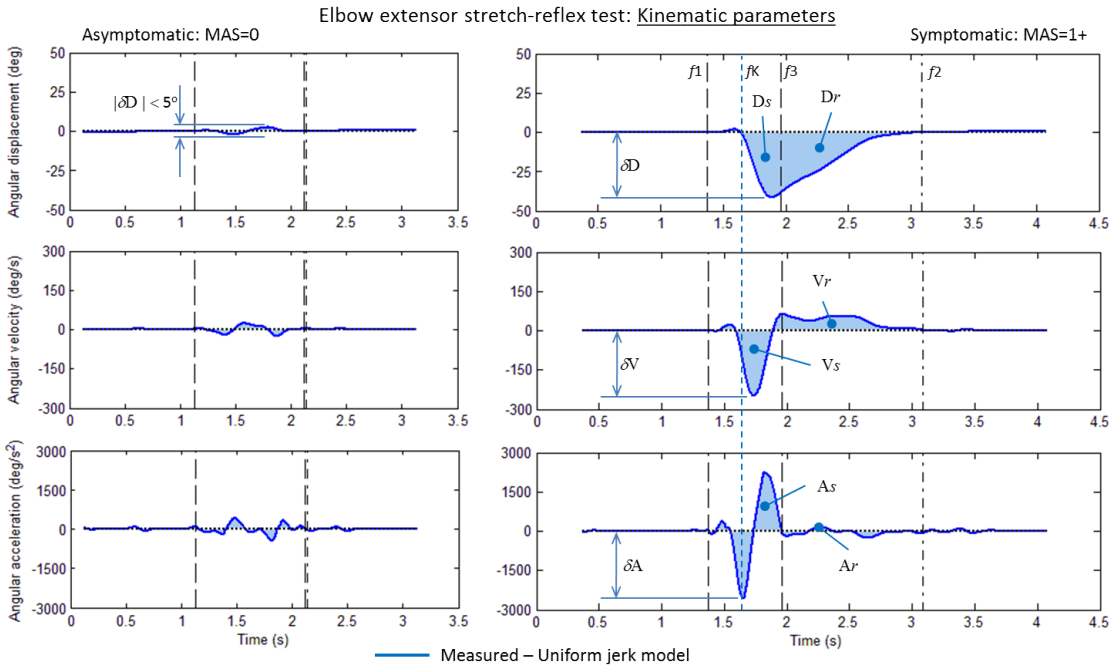

Figure 2. Elbow kinematic departure from uniform jerk model (shaded area); Left panels show departures for asymptomatic extensor with MAS=0; Right panels show departures for symptomatic extensor with MAS=1+. Spasticity onset time (*f*K) was found as the first peak in acceleration (δA) prior to the first peak in velocity (δV) that precedes the peak angular displacement departure (δD). Area under curve was also evaluated for the impulse (D*s*, V*s*, A*s*) and the recovery (D*r*, V*r*, A*r*) periods of the motion. A departure of $\delta D \geq 5^{o}$ was required for onset.

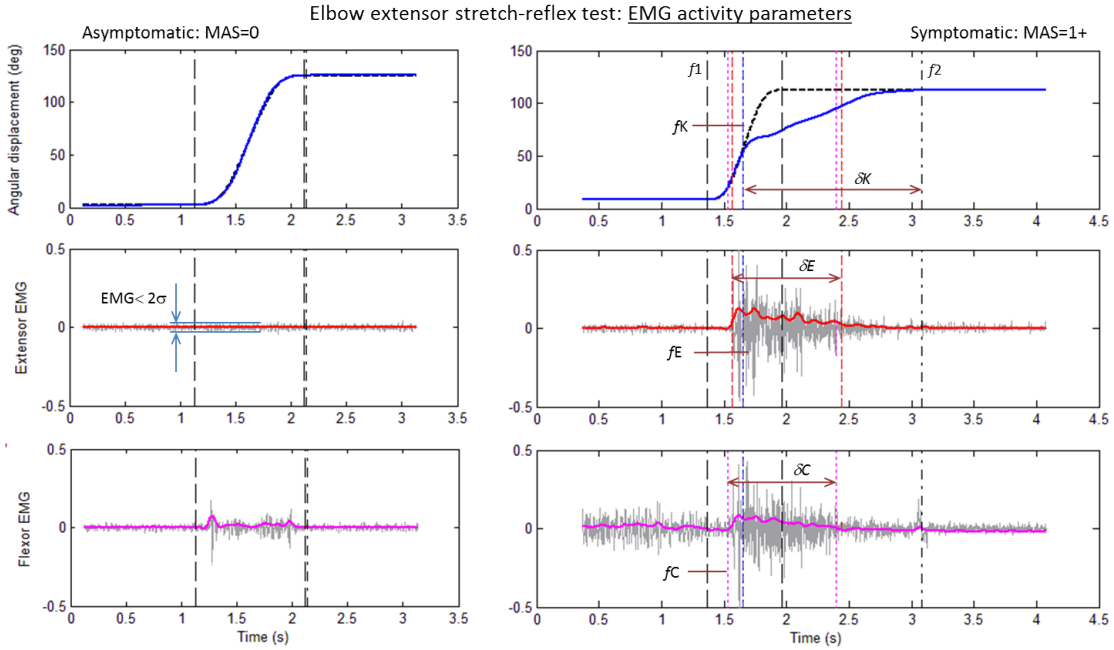

Figure 3. Elbow kinematic profile and predicted spasticity onset time from kinematic data (blue line) and from extensor muscle EMG. The stretched (antagonist) muscle determines spasticity onset (red line). Agonist (co-contraction) activity was also measured (magenta line). EMG onset was time registered (*f*E) when the magnitude exceeded 2 standard deviations from the pre-movement "quiet" EMG signal, and co-contractive muscle (agonist) onset (*f*C) was similarly detected if *f*E was detected.
Time duration of the spastic contraction was measured by duration of the kinematic departure (δK), and EMG activity (δE and δC).

values in angular displacement, velocity and acceleration [20] as well as the area under the curve during the impulse portion (f1 to f3) and the recovery portion (f3 to f2). EMG of biceps (flexors) and triceps (extensors) was acquired simultaneously with kinematic measurement and various metrics extracted as illustrated in Figure 3 and summarized in Table 2. To reduce variability of the SRT, waveform, the measured impulse time (f1-f3) and amplitude (θ_{max}-θ_{min}) was used to scale the measured trajectory to a unit trajectory, as shown in Figure 4. This also enabled ensemble averaging of patient trajectories which was useful for displaying model results.

Table 2. SRT kinematic parameters and elbow strength and ROM measures.

Metric	Description
Flexor and extensor SRT kinematic parameters (normalized to unit trajectory)	
δD, D*s*, D*r*	Angle departure peak, and area under curve for impulse and recovery portions, respectively
δV, V*s*, V*r*	Velocity departure peak, and area under curve for impulse and recovery portions, respectively
δA, A*s*, A*r*	Acceleration departure peak, and area under curve for impulse and recovery portions, respectively
δK	Kinematic departure duration
δE	EMG activity duration of antagonist (stretching) muscle
δC	Co-contraction duration of agonist (shortening) muscle
Flexor and extensor strength & ROM parameters	
Se	Strength of elbow flexors and extensors from LSMD and normalized to body-weight)
Pe	Passive-active limit angle (passive limit angle minus active limit angle in extension or flexion)

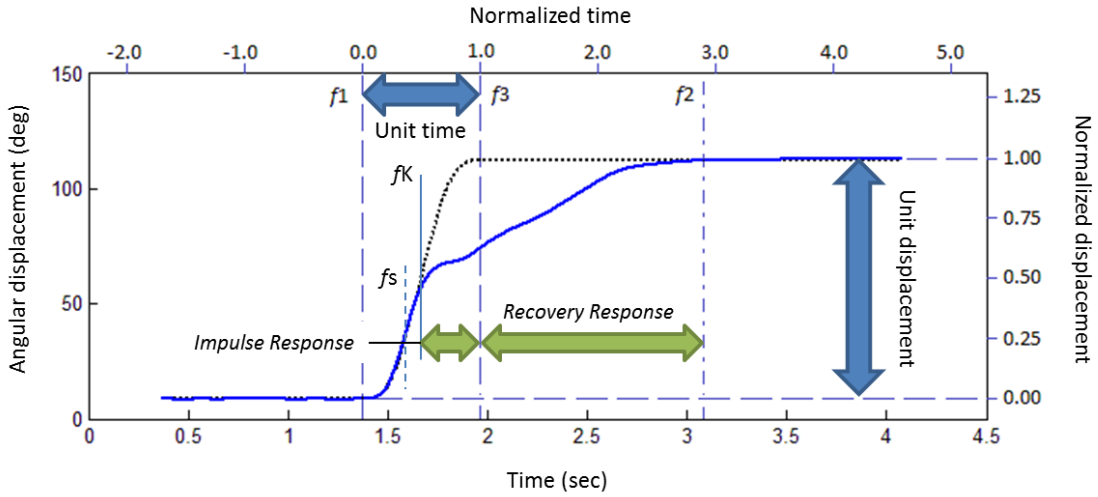

Figure 4. Time and magnitude normalization of stretch-reflex induced kinematics to define the impulse as 1 unit of time to achieve 1 unit of displacement, as shown thick blue arrows, and the response periods for impulse and recovery shown by the green arrows.

Model evaluation

Participants: The BioTone™ system (Figure 5) was deployed at three rehabilitation hospitals from Sept 2011 to May 2014, acquiring data for a total of 103 patients with UMNS due to various etiology. All participants provided informed consent according to multi-institutional ethics approval.

Inclusion criteria were: Male or female active inpatient or outpatient sixteen years of age or older and currently receiving services at study site for one or more of the following diagnoses: acquired brain injury (ABI: strokes, trauma, etc.), spinal cord injury (SCI: incomplete any level or complete C7 and below), multiple sclerosis (MS: meeting MacDonald criteria), and cerebral palsy (CP: hemiplegic or diplegic); medically stable; and exhibits some degree of abnormal tone in either upper or lower limbs, specifically at the elbow and/or knee joint(s).

Exclusion criteria were: Joint arthopathy (osteoarthritis, rheumatoid arthritis, etc.) that would prohibit objective measure of spasticity; bariatric or with little measureable surface EMG signal; viral or bacterial infection; open skin lesions, and; not capable of autonomous consent.

Measurements

Physical therapists and occupational therapists at each site (2-4 per site) were trained to use the BioTone™ system by a research clinician (M.J.) and biomedical engineer (A.S.). Access to an on-line training tool

Figure 5. BioTone™ tools used in the study. A. Fibre-optic goniometer (FOG) was used to measure elbow angle; B. Grip strength measurement device (GSMD); C. Limb strength measurement device (LSMD); D. EMG and FOG set-up for stretch-reflex testing of elbow flexors and extensors.

with professional-grade training videos of the testing procedures was also provided to participating sites. All the BioTone™ test protocols used were previously configured to guide therapists through the testing procedures.

Enrolled patients underwent clinical assessment as regularly scheduled, which included the Modified Ashworth Scale (MAS) assessment of elbow and/or knee flexors and extensors. Age, height, weight, hand dominance, diagnosis (ABI/SCI/MS/CP), affected extremities (left /right arm and/or leg), and year and month of injury or onset, were also recorded. Only data and procedures for the elbow are discussed here.

BioTone™ tools were then used to acquire measurements of grip strength, elbow flexor and extensor [20]. BioTone™ sensors had a sampling rate of 1000Hz. Measures used in the analysis are summarized in Table 2.

Strength measurement: Custom devices for grip strength (GSMD), and elbow flexor and extensor strength (LSMD) [26] were used as shown in Figures 5b and 5c, respectively. The GSMD was specifically designed to measure grip strength of hands with limited dexterity. The LSMD has a fixed moment arm, can be configured to measure elbow extensor strength and flexor strength, slides onto the arm to accommodate contracture, and is easily removed if spasms occur. Patients were instructed to "push their forearm (or squeeze hand) as hard as possible". Maximal grip and elbow flexor and extensor forces were each measured and automatically recorded three times in succession. Mean values were used in the data analysis after normalizing to the patient's body-weight in Newtons.

ROM testing: Participants donned a low-profile fibre-optic goniometer (FOG, Figure 5a) (ShapeSensor™, Measurand Inc. NB) in an upright seated position during passive and active ROM testing. For passive ROM the therapist moved the limb through its full, unrestricted range where the minimum (extension) and maximum (flexion) angles were recorded. For active ROM the patient was instructed to move their own limb through its full range, where again the minimum (extension) and maximum (flexion) angles were recorded. Tests were repeated three times each. Passive-active limit angles were then calculated as the mean passive minus mean active range limits (i.e., in flexion and extension).

Stretch-reflex testing: During the SRT (Figure 5d), the FOG was used in combination with a custom 2-channel EMG system (Ag/AgCL duotrode electrode) during a manual SRT using the same testing protocol as used for the MAS grading prior to BioTone™ measurements. EMG was recorded from biceps and triceps during all tests. SRTs were performed three times on elbow extensors (fast passive flexion) and three times for elbow flexors (fast passive extension). Only the rapid trials are analyzed here. SRT parameters (described above and listed in Table 2) were then calculated and averaged across the repeated tests for extensors and flexors separately.

Model validation and translation

Validation of the model framework was tested by: 1) Evaluating the correlation and time-delay between the kinematic model's prediction of spasticity onset and duration, with measured EMG onset time and duration, during the SRT, and; 2) Evaluating the bivariate correlation between BioTone™ SRT metrics (kinematic and EMG) with therapist-rated MAS scores measured prior to BioTone™ assessments.

Parameters in Table 2 were entered into a linear discriminant analysis (LDA) as predictors and the therapist rated MAS score for the corresponding muscle (flexor or extensor) was entered as the criterion variable. Predictors were entered all at once with prior probabilities computed from group size and within-groups covariance matrix. Participants were classified into predicted MAS categories (BT-MAS) and a cross-tabulation analysis was then conducted to evaluate the classification accuracy.

Probability of membership in each MAS category was then used as weighting factors of the weighted sum of the five indexed categories, which produces a continuous scale score that spans the 0-4 range of MAS scoring. Due to the complication of the "1+" category, it was decided not to use a 0-4 scale for the BT-MAS score, but rather to use a 0-10 scale with precision of .1 unit (101 points). Regression analysis R^2 values were used to evaluate the extent to which the BT-MAS could explain therapists' MAS ratings. A custom report (in Matlab) was designed to present the BT-MAS score as a mapping of predicted score onto therapist-rated MAS scores.

SPSS Version 20 (SPSS-PAS, IBM Corp.) was used for all statistical analysis, with significance level of α=.05.

Results

Participant demographics and clinical presentation by diagnostic cohort is shown in Table 3. Of the 103 participants in the larger study, 79 were examined unilaterally (most involved side) for upper- extremity involvement consisting of ABI (n=53), SCI (n=12), CP (n=8), and MS (n=6). Fifty-three of the 79 participants were male and age ranged from 16 to 93 years ($\bar{x}$=52.2 ± 16.9 year). Elbow flexor and extensor MAS scores spanned the scale from 0 to 3 and varied by diagnostic group, where ABI and CP had the highest level of spasticity (1+ to 3) while MS and SCI had lower levels of spasticity (0 to 1+).

BioTone™ measures of passive and active ROM and upper extremity strength measures are summarized by diagnostic cohort in Table 4. ABI patients had the lowest passive ROM and correspondingly the lowest active ROM, and SCI patients the highest passive ROM and active ROM. Strength measures were quite low for sample as a whole. Although strength measures were normalized by body weight for the data analysis that follows, the parameters shown in Table 4 are in

Table 3. Patient demographics and Modified Ashworth Scale (MAS) scores of elbow spasticity. The table shows means and 95% confidence intervals.

Dx	Gender			Age, years		Time since Dx., years		MAS* Flexors		MAS* Extensor	
	M	F	Total	Mean	(CI^{95})	Mean	(CI^{95})	Mean	(CI^{95})	Mean	(CI^{95})
	n	*n*	*n*								
ABI	39	14	53	55.1	(50.7-59.5)	4.5	(3.2-5.8)	2.0	(1.7-2.3)	1.6	(1.3-1.9)
MS	2	4	6	59.3	(50.7-67.9)	15.4	(8.1-22.7)	1.0	(-0.4-2.4)	.2	(-0.1-0.5)
CP	3	5	8	30.6	(23.3-37.9)	30.6	(23.3-37.9)	2.4	(1.6-3.2)	1.9	(1.1-2.7)
SCI	9	3	12	50.3	(41.6-59.0)	3.6	(-1.0-8.2)	.9	(0.3-1.5)	.8	(0.2-1.4)
Total	53	26	79	52.2	(48.5-55.9)	8.1	(5.8-10.4)	1.8	(1.5-2.1)	1.4	(1.1-1.7)

Dx = Diagnostic cohort: ABI = Acquired brain injury (stroke, trauma); MS = Multiple sclerosis; CP = Cerebral palsy; SCI = Spinal cord injury. Numeric MAS used for averaging (0="0", 1="1", 2="1+", 3="2",4="3"). CI^{95} = 95% confidence interval (lower bound – upper bound).

Table 4. BioTone™ measures for upper-extremity passive and active ROM, grip strength and elbow flexor and extensor strength. The table shows means and 95% confidence intervals.

Dx	Range of Motion, deg				Strength, N					
	Passive		Active		Grip		Elbow Flexor		Elbow Extensor	
	Mean	(CI^{95})	Mean	(CI^{95})	Mean	(CI^{95})	Mean	(CI^{95})	Mean	(CI^{95})
ABI	121.5	(114.3-128.7)	86.1	(76.0-96.3)	86.4	(68.1-104.7)	63.8	(48.9-78.6)	102.1	(84.2-120.0)
MS	134.8	(111.8-157.7)	101.3	(56.1-146.6)	143.4	(45.8-240.9)	85.7	(34.2-137.2)	122.0	(51.5-192.5)
CP	131.3	(118.7-143.9)	106.7	(94.0-119.5)	124.7	(66.8-182.7)	78.3	(45.0-111.6)	81.9	(54.8-109.1)
SCI	133.5	(126.0-141.1)	121.5	(113.5-129.5)	152.9	(19.8-286.0)	83.0	(40.9-125.1)	92.6	(28.8-156.4)
Total	124.9	(119.3-130.5)	93.8	(85.4-102.2)	101.2	(79.1-123.4)	69.1	(56.3-81.9)	100.6	(84.5-116.6)

Dx = Diagnostic cohort: ABI = Acquired brain injury (stroke, trauma); MS = Multiple sclerosis; CP = Cerebral palsy; SCI = Spinal cord injury. CI^{95} = 95% confidence interval (lower bound – upper bound).

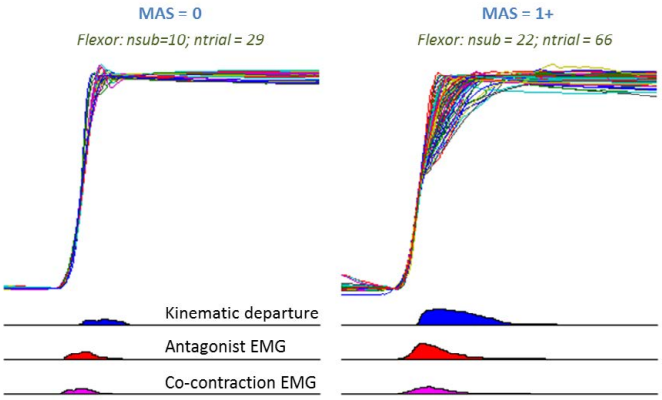

Figure 6. Examples of time and range normalized signals for generating the model parameters. Left: Asymptomatic flexors (MAS=0) for 29 trials of 10 subjects; Right: Symptomatic (MAS=1+) flexors for 66 trials of 22 subjects. Bottom: Histograms of spastic duration for kinematic (blue), antagonist EMG (red) and co-contraction (magenta).

Newtons. Grip strength and elbow flexor strength was lowest in the ABI group. MS patients had the most elbow flexor and extensor strength, and CP and SCI patients had equally low flexor and extensor strength, but SCI patients had better grip strength than the other groups.

Sixty-three of the participants had complete kinematic and EMG records for both flexor and extensor testing. As the goal was to address spasticity assessment across the spectrum of UMNS (including age, gender and circumstance) the results focus on pooled sample of n=63.

Validation of the model

Figure 6 shows a compilation of normalized angle-time histories for low tone (MAS=0) elbows and moderate tone (MAS=1+) and corresponding histograms of spastic duration from kinematic model and EMG activity. Muscle onset was detected in 81% of patients with symptoms (MAS>0), and no onset was detected in 75% of patients with MAS=0. Chi-square analysis showed a significant relationship between onset detection and symptoms of spasticity (χ^2(df=1)=35.0, p<.001).

In order to validate the kinematic model parameters as reflecting physical manifestation of muscle activation, first the kinematic onset time fK was compared to the antagonist EMG onset time fE. Figure 7a shows the strong correlation between the kinematic and EMG onset time estimates; fK explained approximately 70% of the variance in fE. Correlation between EMG measured duration of muscle activity (δE) and kinematic estimation of spasticity duration (δK) was very high for both flexors (ρ = .91, p<.001) and extensors (ρ=.96, p<.001). For those with muscle onset detected, a within-subjects ANOVA revealed that the difference between fK and fE was significant for muscles (F(df=1,97)=239.8, p<.001) with a mean of 120 ms (95% CI: 136-105 ms), but was not different between muscles (interaction,

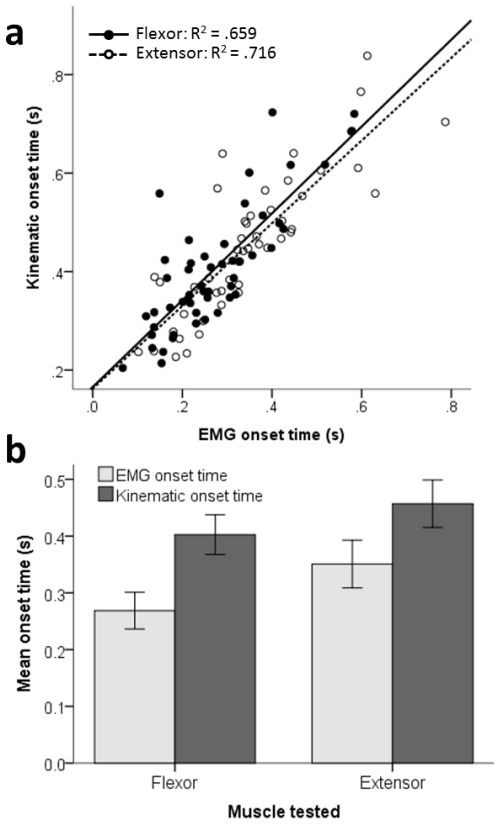

Figure 7. Spasticity onset prediction from the kinematic model and EMG. Top: The scatter plot between EMG and kinematic onset predictions demonstrates a high degree of correlation, strongly suggesting they are measuring the same phenomena. Bottom: Bar charts with 95% confidence intervals on EMG measured (light) and kinematic predicted (dark) onset time. Not surprisingly the EMG onset preceded the external torque response of therapist by ~120ms which is consistent with expected delays in haptic feedback.

F(df=1,97)=3.19, p=.077). Onset time results are shown as the bar chart in Figure 7b.

Next, normalized kinematic and EMG parameters extracted from the model were explored for their relevance to therapist-rated muscle spasticity. Top and middle plots of Figure 8 shows bar charts of means and 95% confidence intervals of model parameters derived from angular displacement departure, across MAS categories. Time normalized antagonist EMG activity duration and kinematic departure duration are shown in the bottom of Figure 8.

Spearman correlation between model parameters and MAS scores for flexors and extensors separately are shown in Table 5. Correlation coefficients were slightly stronger for extensors compared to flexors, but almost all of the derived parameters had strong positive correlations

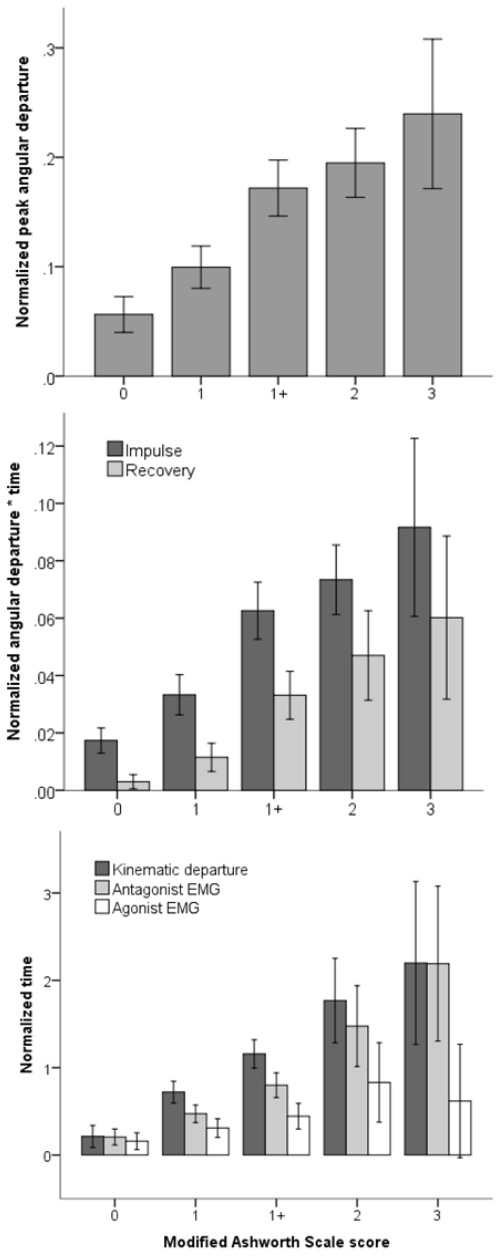

Figure 8. Kinematic metrics by therapist-rated MAS score (pooled 126 muscles): Bars are mean normalized metrics with 95% confidence limits on the MAS category mean. Top: Normalized peak angular displacement; Middle: Normalized departure "density" (area under curve) or angular displacement * time. Bottom: Kinematic and EMG metrics of spastic response duration (normalized time) by therapist-rated MAS score, where EMG duration is shown for both the stretched antagonist muscle (muscle being tested) and the shortened agonist represents the contracting muscle. Across all metrics the variability was higher for MAS scores 2 and 3, nevertheless the metric increase linearly in magnitude with therapist-rated MAS scores.

with therapist-rated MAS score. Normalized angular departure metrics (δD, Ds, Dr) and spasticity duration (δE, δK) had the highest correlations, ranging between ρ=.63 and .79 (all significant at p<.001), followed by velocity parameters (δV, Vs, Vr) with correlations between ρ=.43 and .69 (p<.001). Correlations were lower for acceleration parameters, especially for recovery phase kinematics which were non-significant. There were weak to non-significant correlations with duration of agonist (co-contractive) EMG.

Translation of the model results

The first step was transforming the set of metrics into MAS categories using a supervised LDA. The LDA produced classification accuracy of 74% for elbow extensors and 71% for elbow flexors. However, 93% and 95% of cases, respectively, were classified within +/- 1 MAS category, which is not surprising given that MAS ratings are subjective (ie. not an ideal "gold standard"). Interestingly, a drop-one-out analysis yielded poor classification accuracy (<50%) indicating the robustness of the LDA likely suffered from insufficient power, given the moderately high variability of model parameters within higher MAS score groups. The cross-tabulation results are shown in Table 6.

Then the probability scores from the LDA were used as described above to generate the BT-MAS score on a 0-10 scale. The derived score correlated significantly with the therapist-rated MAS scores, explaining almost 76% of the variance, and the predicted MAS score, explaining 92% of the variance. These data are shown in Figure 9. This allowed the display of any participant's BT-MAS score mapped onto the probability of belonging to each MAS category, presented as a color coded visual analog. Figure 10a shows an example of a clear consensus between therapist-rated score and BT-MAS score for flexor muscle, but incorrectly classified for extensor in Figure 10b. This illustrates that misclassification was often characterized by a test result that straddled the boundary between adjacent MAS categories.

Discussion

The need for objective and sensitive clinical assessment of spasticity is made clear by the number of studies that have questioned the use of clinical rating scales [8,9,27,28] and their validity as a measure of spasticity [10,11,29]. Although the latter studies used motorized solutions for objective and precise testing, there are barriers to their adoption into the clinic [30]. EMG [19,20,31] and kinematic sensors [19,20,21] are more feasible for objective measurement in the clinic, but without a framework for analyzing and translating this data the added value of this solution may too not be realized, and thus not get adopted by practitioners and researchers who need it. The motivation of this work was to test a conceptually simple model for analyzing the sensor signals during the manual SRT and translating results into information that can be used by the therapist. The much maligned but ubiquitously employed MAS was explored in this work to better understand the SRT response from a neuro-biomechanical perspective and to demonstrate the feasibility and added value of a wearable sensor system during a commonly used elbow SRT protocol in the clinic.

Although a primary goal was to address the need for more objective and sensitive measures of spasticity, the proposed approach could also enable the development of a large repository of spasticity assessments, made possible by the normalization scheme (Figure 4), and the supporting infrastructure of the BioTone™ system. Such a system could have significant benefits for conducting multi-site clinical trials that include elbow MAS as a primary outcome.

Translation of instrumented muscle tone assessment

For wearable sensors to be of value to therapist and other end-users the information they generate must be translatable, conform to current reporting practices, and add considerable value to what already can be done. We demonstrated using a supervised LDA that the probability of MAS category membership for each patient's SRT of flexors or extensors could be used to assign scores on much refined scale.

Figure 10 shows examples that help demonstrate how such a

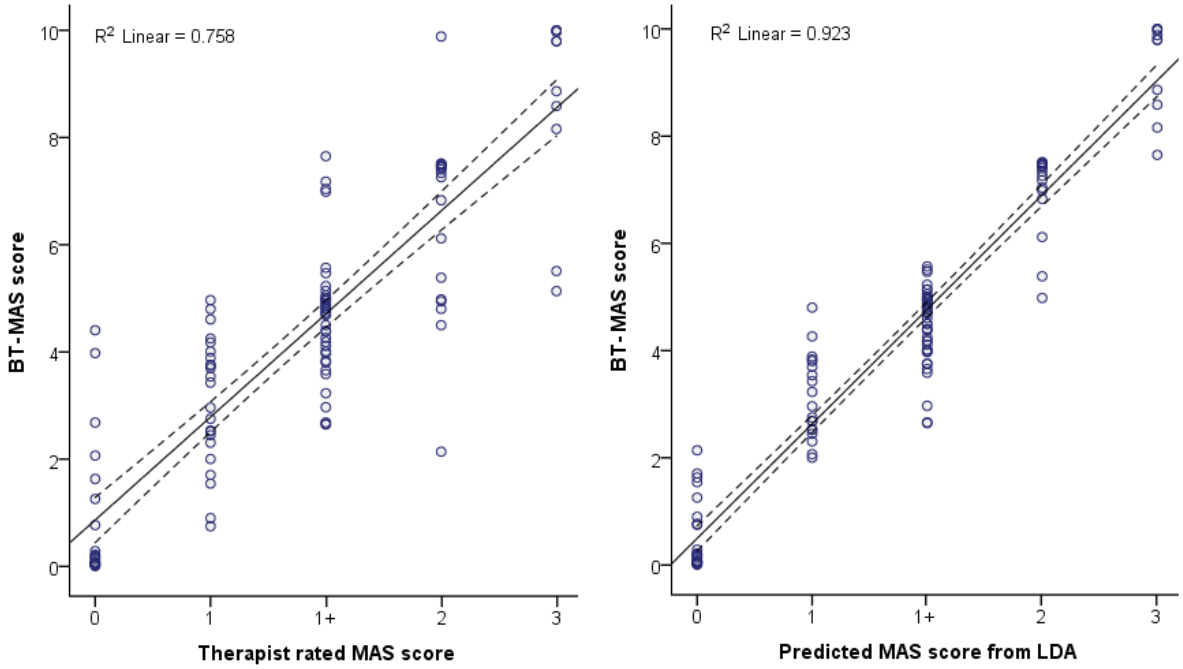

Figure 9. Probability-based BT-MAS metric versus therapist-rated (left) and LDA predicted (right) MAS score for all muscles tested (n=126). The horizontal axis was recoded for display purposes (internally the MAS score represented as an integer from 0 to 4). Because of the 1+ category, coding the continuous BT-MAS between 0 and 3 could lead to confusion, so the new metric was scaled from 0-10.

Table 5. Spearman correlations between model output parameters and MAS score. There were a total of n=63 patients with complete SRT records (joint kinematics and EMG).

A. Flexor MAS	δD	δV	δA	*Ds*	*Vs*	*As*	*Dr*	*Vr*	*Ar*	δE	δC	δK
Spearman ρ	.626	.524	.303	.670	.468	.325	.673	.425	.012	.792	.184	.690
p-value	<.001	<.001	.016	<.001	<.001	.009	<.001	.001	.926	<.001	.149	<.001
N	63	63	63	63	63	63	63	63	63	63	63	63
B. Extensor MAS	δD	δV	δA	*Ds*	*Vs*	*As*	*Dr*	*Vr*	*Ar*	δE	δC	δK
Spearman ρ	.724	.666	.451	.765	.685	.536	.690	.523	.294	.700	.379	.721
p-value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.019	<.001	.002	<.001
N	63	63	63	63	63	63	63	63	63	63	63	63

metric can be presented in a way that can be interpreted. Figure 10a shows the calculated BT-MAS score for a muscle where there was agreement between therapist-rated and LDA predicted MAS score (92% probability); both rated the flexor muscle spasticity at MAS=1+. The corresponding BT-MAS score was 4.8/10. Figure 10b shows the result for the same patient's elbow extensor, showing this time a disagreement between therapist-rated MAS=1+ and the LDA predicted MAS=2, reporting a 83% probability of being correct, but also a 16% probability the extensor was indeed MAS=1+.

The example in Figure 10b illustrates the added value of using wearable sensors during routine SRT with the proposed framework and scale generation; patients whose muscles are straddling the transition between MAS categories could have a more precise rating and potentially avoid inappropriate treatment prescription. Given the tremendous economic burden being placed on the health care system due to the high prevalence of stroke and the life-long effects in younger adults with spinal cord and brain injury, the ability to assess muscle spasticity in the clinic with a higher degree of objectively and more precise level of measurement is greatly needed. The work presented here provides hope this goal may be within reach.

Although we focused on explaining the MAS rating, there are other clinical rating tools available to clinicians. The Modified Tardieu scale (MTS) is based on a slow and fast SRT, where the slow stretch motion is used to define the passive range, and fast trial is used to define the catch angle. The difference between these two angles is reported to be a measure of the hyper-excitability of the stretch-reflex and therefore is better differentiated from contracture [32]. However, this test also suffers from reliability and repeatability issues [27,28], particularly due to the need to reposition the limb at the catch angle to acquire the measurement. However, better repeatability has been achieved using inertial sensors [21]. Although we do not report Tardieu scores here, the SRT protocol we used and the analysis (detection of the onset time from kinematics) could also be used to quantify Tardieu index scores (catch angle / passive range).

Our study also supports the notion that the clinical MAS rates two phenomena simultaneously. The SRT may induce two components of spastic muscle contraction: the catch/release, a highly prominent feature of MAS score categories 1 and 1+ where the muscle suddenly contracts and relaxes during passive motion, and; the progressive resistance to passive movement that becomes a more prominent feature of MAS score categories 2 and 3, where the muscle stays active for a prolonged duration. The data in Figure 8 suggest that both of these characteristics, as indicated by the peak departure (catch) and departure duration (prolonged resistance) increase in a linear fashion with increasing therapist-rated MAS score. Although peak departure

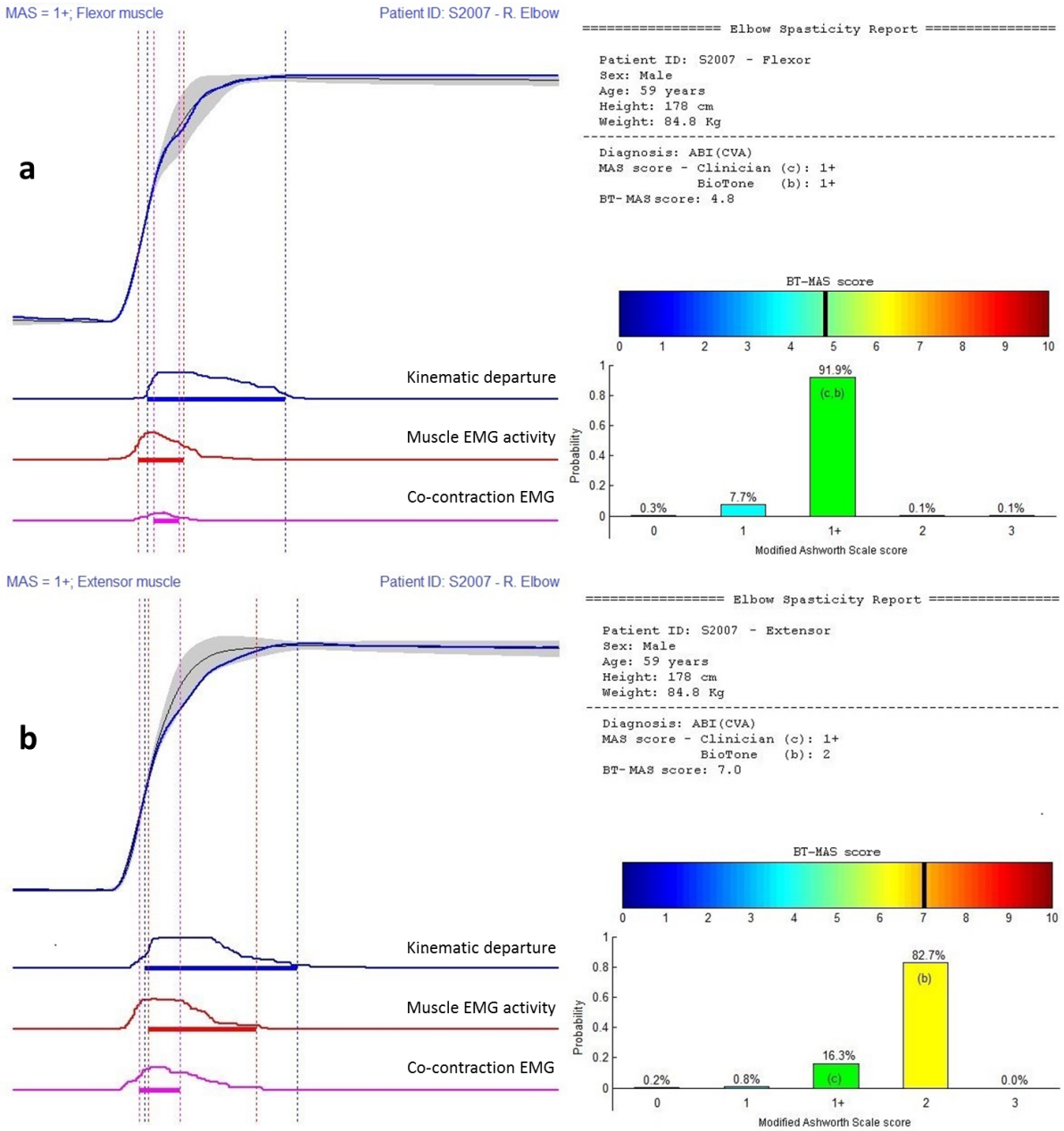

Figure 10. A representative patient with symptomatic flexors and extensors, where the therapist rated both muscles as MAS=1+, which was correctly classified for flexors (a), but for extensor the predicted category was MAS=2 (b). Plots on the left show the patient's normalized angular profile against all others assigned to the same category with a correct classification; each also shows the kinematic departure and EMG durations (solid lines) against the group histogram. The right side shows a mock report describing the patient (59 y.o. male with stroke) and the calculated BT-MAS score; it can be seen that the extensor muscle had more spasticity than the flexor muscle, but to the therapist had similar presentation.

effects appear to tail-off at higher MAS categories, there were too few patients in the MAS=3 category to make any firm conclusions about this behavior. Nevertheless, this finding is consistent with observations that the MAS may be confounded by contracture and becomes a less reliable tool for clinical assessment in patients with high tone [23].

Limitations and outstanding issues

Several important limitations of the study are worthy of discussion. One of these limitations is that we only examined elbow flexor and extensor spasticity and measured only EMG from biceps brachii and triceps brachii. Stroke and higher-level spinal cord injury often affects brachialis and brachioradialis as well as muscles of the shoulder, wrist and hand. Although it is unknown if the kinematic profile of passive wrist (or ankle, etc.) SRT would obey the uniform jerk assumption, it is possible the proposed approach may be applicable to tone assessment of other joint muscles.

A significant limitation is that there is no true gold standard by which to conduct a robust classification of spasticity using the framework metrics. Therapists participating in the study were trained to use the BioTone™ system and all therapists were familiar and qualified to administer the MAS. A thorough review of every muscle testing record observed very few trials where the kinematic profile

Table 6. The numbers in the diagonal cells (dark gray) cannot be seen and surely will not reproduce well. Therefore I suggest the numbers in these cells be white text (as originally submitted) or the cell color be changed to a lighter shade.

Flexors		Predicted MAS score from LDA					Total
		0	1	1+	2	3	
Therapist-rated MAS score	0	8	0	0	0	0	8
	1	2	7	4	0	0	13
	1+	0	3	15	2	1	21
	2	0	1	1	5	1	8
	3	0	0	2	0	7	9
Total		10	11	22	7	9	59
		Exact classification					71.2%
		+/- 1 category					93.2%
		Drop one-out analysis					35.6%
Extensors		Predicted MAS score from LDA					Total
		0	1	1+	2	3	
Therapist-rated MAS score	0	14	2	2	0	0	18
	1	2	4	4	0	0	10
	1+	0	2	14	1	0	17
	2	1	0	1	10	0	12
	3	0	0	0	0	1	1
Total		17	8	21	11	1	58
		Exact classification					74.1%
		+/- 1 category					94.8%
		Drop one-out analysis					46.6%

suggested a deviation from protocol, and indeed the LDA provided some convincing evidence that some MAS scores (~20%) could have been off by +/- 1 category.

Another factor potentially causing misclassification was if the therapist "drove through" the spasticity, resulting in little kinematic departure and subsequently misclassified as a lower MAS score than what the therapist-rated. It is important to note that therapists were only instructed to use the standard SRT protocol used for the MAS (Bohannon and Smith 1987). Furthermore, MAS rating were performed as part of the patient's routine clinical examination – the instrumented SRTs were conducted after the regular examination with other BioTone™ measurements. In this light, it was encouraging that 126 muscle tests at three different rehab hospitals by several different therapists on four different patient populations resulted in such uniform and sensible results.

Although complete records were available for 126 elbow muscle SRTs (63 elbows), the patient distribution was not uniform across MAS categories, with 27% MAS=0, 18% MAS=1, 28% MAS=1+, 18% MAS=2 and 9% MAS=3. Future studies are needed to increase sample sizes and distribution across categories. Larger samples would allow study of other factors, such as gender, age, time since injury, and other factors relevant to managing spasticity in UMNS patients.

Finally, although we used a LDA to classify patients by their probability of belonging to a MAS category, there are many choices available for such a task. Support vector machines and hidden Markov models are other examples whereby a classifier could be trained to identify and assign objective spasticity scores. Although there are clearly limitations to machine learning methods when the quality of gold standard used is itself questionable, the approach shows some promise for delivering objective spasticity measurement to research and practice.

Authorship and contribution

All authors contributed significantly to the work. McGibbon was the PI of the study and O'Connell the co-I, who jointly conceived and planned the research. Sexton was the biomedical engineer that developed the BioTone™ system and managed the site studies, and Jones was the research therapist that co-designed the protocols and trained the therapist to the use the BioTone™ system. The biomechanical and statistical analyses were conducted by McGibbon. All authors have read and approve of the submitted manuscript.

Acknowledgments

The authors are indebted to contributing staff and graduate students of the UNB Institute of Biomedical Engineering and Stan Cassidy Centre for Rehabilitation, Nova Scotia Rehabilitation Hospital in Halifax, and Spaulding Rehabilitation Hospital in Boston.

Funding information

This work was supported by the Atlantic Canada Opportunities Agency, Atlantic Innovation Fund, Project # 195180.

Competing interest

The authors declare that they have no competing interests.

References

1. Damage to descending motor pathways: The upper motor neuron syndrome (2001) In: Purves D, Augustine GJ, Fitzpatrick D (Eds.), Neuroscience. (2ndedn), Sunderland (MA): Sinauer Associates.
2. Dromerick AW (2002) Clinical features of spasticity and principles of treatment. In: Clinical evaluation and management of spasticity. Gelber DA, Jeffry DR (Eds.), Totowa, NJ: Humana Press, 13.
3. Shackley P, Shaw L, Price C, van Wijck F, Barnes M, et al. (2012) Cost-effectiveness of treating upper limb spasticity due to stroke with botulinum toxin type A: results from the botulinum toxin for the upper limb after stroke (BoTULS) trial. *Toxins (Basel)* 4: 1415-1426. [Crossref]
4. Chang E, Ghosh N, Yanni D, Lee S, Alexandru D, et al. (2013) A Review of Spasticity Treatments: Pharmacological and Interventional Approaches. *Crit Rev Phys Rehabil Med* 25: 11-22. [Crossref]
5. Malhotra S, Pandyan AD, Day CR, Jones PW, Hermens H (2009) Spasticity, an impairment that is poorly defined and poorly measured. *Clin Rehabil* 23: 651-658. [Crossref]
6. Bohannon RW, Smith MB (1987) Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther* 67: 206-207. [Crossref]
7. Lance JW (1980) Symposium synopsis. In: Spasticity: Disordered motor control. Feldman RG, Young RR, Koella WP (Eds.), Chicago: Year book Medical Publishers, 485.
8. Ansari NN, Naghdi S, Moammeri H, Jalaie S (2006) Ashworth Scales are unreliable for the assessment of muscle spasticity. *Physiother Theory Pract* 22: 119-125. [Crossref]
9. Fleuren JF, Voerman GE, Erren-Wolters CV, Snoek GJ, Rietman JS, et al. (2010) Stop using the Ashworth Scale for the assessment of spasticity. *J Neurol Neurosurg Psychiatry* 81: 46-52. [Crossref]
10. Alibiglou L, Rymer WZ, Harvey RL, Mirbagheri MM (2008) The relation between Ashworth scores and neuromechanical measurements of spasticity following stroke. *J Neuroeng Rehabil* 5: 18. [Crossref]
11. Kumar RT, Pandyan AD, Sharma AK (2006) Biomechanical measurement of post-stroke spasticity. *Age Ageing* 35: 371-375. [Crossref]
12. Malhotra S, Cousins E, Ward A, Day C, Jones P, et al. (2008) An investigation into the agreement between clinical, biomechanical and neurophysiological measures of spasticity. *Clin Rehabil* 22: 1105-1115. [Crossref]
13. Damiano DL, Quinlivan JM, Owen BF, Payne P, Nelson KC, et al. (2002) What does the Ashworth scale really measure and are instrumented measures more valid and precise? *Dev Med Child Neurol* 44: 112-118. [Crossref]
14. de Vlugt E, de Groot JH, Schenkeveld KE, Arendzen JH, van der Helm FC, et al. (2010) The relation between neuromechanical parameters and Ashworth score in stroke patients. *J Neuroeng Rehabil* 7: 35. [Crossref]

15. Pierce SR, Johnston TE, Shewokis PA, Lauer RT (2008) Examination of spasticity of the knee flexors and knee extensors using isokinetic dynamometry with electromyography and clinical scales in children with spinal cord injury. *J Spinal Cord Med* 31: 208-214. [Crossref]

16. Pisano F, Miscio G, Del Conte C, Pianca D, Candeloro E, et al. (2000) Quantitative measures of spasticity in post-stroke patients. *Clin Neurophysiol* 111: 1015-1022. [Crossref]

17. Pizzi A, Carlucci G, Falsini C, Verdesca S, Grippo A (2005) Evaluation of upper-limb spasticity after stroke: A clinical and neurophysiologic study. *Arch Phys Med Rehabil* 86: 410-415. [Crossref]

18. Naghdi S, Ansari NN, Mansouri K, Asgari A, Olyaei GR, et al. (2008) Neurophysiological examination of the Modified Modified Ashworth Scale (MMAS) in patients with wrist flexor spasticity after stroke. *Electromyogr Clin Neurophysiol* 48: 35-41. [Crossref]

19. Calota A, Feldman AG, Levin MF (2008) Spasticity measurement based on tonic stretch reflex threshold in stroke using a portable device. *Clin Neurophysiol* 119: 2329-2337. [Crossref]

20. McGibbon CA, Sexton A, Jones M, O'Connell C (2013) Elbow spasticity during passive stretch-reflex: clinical evaluation using a wearable sensor system. *J Neuroeng Rehabil* 10: 61. [Crossref]

21. Paulis WD, Horemans HL, Brouwer BS, Stam HJ (2011) Excellent test-retest and inter-rater reliability for Tardieu Scale measurements with inertial sensors in elbow flexors of stroke patients. *Gait Posture* 33: 185-189. [Crossref]

22. Jobin A, Levin MF (2000) Regulation of stretch reflex threshold in elbow flexors in children with cerebral palsy: a new measure of spasticity. *Dev Med Child Neurol* 42: 531-540. [Crossref]

23. Pandyan AD, Van Wijck FM, Stark S, Vuadens P, Johnson GR, et al. (2006) The construct validity of a spasticity measurement device for clinical practice: an alternative to the Ashworth scales. *Disabil Rehabil* 28: 579-585. [Crossref]

24. Flash T, Hogan N (1985) The coordination of arm movements: an experimentally confirmed mathematical model. *J Neurosci* 5: 1688-1703. [Crossref]

25. Kistemaker DA, Wong JD, Gribble PL (2014) The cost of moving optimally: kinematic path selection. *J Neurophysiol* 112: 1815-1824. [Crossref]

26. Landry J, Sexton A, Hughes G, McGibbon CA (2015) inventors; May 12. Limb strength measurement device. 9,028,433.

27. Ansari NN, Naghdi S, Hasson S, Azarsa MH, Azarnia S (2008) The Modified Tardieu Scale for the measurement of elbow flexor spasticity in adult patients with hemiplegia. *Brain Inj* 22: 1007-1012. [Crossref]

28. Mackey AH, Walt SE, Lobb G, Stott NS (2004) Intraobserver reliability of the modified Tardieu scale in the upper limb of children with hemiplegia. *Dev Med Child Neurol* 46: 267-272. [Crossref]

29. Lindberg PG, Gäverth J, Islam M, Fagergren A, Borg J, et al. (2011) Validation of a new biomechanical model to measure muscle tone in spastic muscles. *Neurorehabil Neural Repair* 25: 617-625. [Crossref]

30. Marinelli L, Trompetto C, Mori L, Vigo G, Traverso E, et al. (2013) Manual linear movements to assess spasticity in a clinical setting. *PLoS One* 8: e53627. [Crossref]

31. Cooper A, Musa IM, van Deursen R, Wiles CM (2005) Electromyography characterization of stretch responses in hemiparetic stroke patients and their relationship with the modified ashworth scale. *Clin Rehabil* 19: 760-6. [Crossref]

32. Patrick E, Ada L (2006) The Tardieu Scale differentiates contracture from spasticity whereas the Ashworth Scale is confounded by it. *Clin Rehabil* 20: 173-182. [Crossref]