Identification of abnormal motor cortex activation patterns in children with cerebral palsy by functional near-infrared spectroscopy

Bilal Khan†
Fenghua Tian‡
Khosrow Behbehani
Mario I. Romero
University of Texas at Arlington
and
University of Texas Southwestern Medical Center at Dallas
Joint Graduate Program in Biomedical Engineering
Arlington, Texas 76019

Mauricio R. Delgado
Nancy J. Clegg
Linsley Smith
Dahlia Reid
Texas Scottish Rite Hospital for Children
Department of Neurology
2222 Welborn Street
Dallas, Texas 75219

Hanli Liu
George Alexandrakis†
University of Texas at Arlington
and
University of Texas Southwestern Medical Center at Dallas
Joint Graduate Program in Biomedical Engineering
Arlington, Texas 76019

1 Introduction

Cerebral palsy (CP) occurs in 2 per 1000 live births every year. CP describes a set of disorders that prevent proper movement and posture often accompanied by loss of sensation, perception, cognition, communication, secondary musculoskeletal problems, and recurrent seizures. This disorder is caused by damage to the motor control centers of the developing brain. The damage can occur during pregnancy, childbirth, or after birth early in a child’s life. One of the most prevalent types of the disorder is hemiparetic CP, an incomplete paralysis of one-half of the body. These motor deficits profoundly affect a child’s ability to develop motor skills and to engage fully in play, exploration, and self-help activities. Thus, children with CP spontaneously undergo plastic changes in their neural circuitry by recruiting healthy portions of their brain to partially compensate for the affected areas, resulting in their improved motor performance. Clinical professionals currently diagnose the functional severity of a child with CP based on qualitative scores derived by observing the success with which that child handles a prescribed series of upper extremity tasks. However, no standard imaging method currently exists for observing the spatiotemporal activation patterns of the motor cortex in children with CP and how these patterns differ from patterns of healthy children.

Although functional magnetic resonance imaging (fMRI) is the “gold standard” for neuroimaging, it often requires the patient’s complete body confinement, steadiness, and minimal interference caused by motion artifacts for the duration of experiments, which limits the success rate to <50% for normal children. Nevertheless, there are case reports on the use of fMRI alone or in combination with transcranial magnetic stimulation, which investigated different aspects of cortical reorganization in children with CP. Although CP is thought to originate mostly from subcortical lesions and less often from cortical ones, there is evidence that associated plasticity changes occur in the motor cortex. Certain differences in cortical activation patterns have emerged from these studies between healthy and children with CP. For example, in an fMRI study, children with hemiparesis of their upper extremities showed stronger ipsilateral activation when tapping with their affected hand. In the future, identifying such differential activation patterns could potentially provide biomarkers to help physicians assess the severity of CP in a more quantita-

Abstract. We demonstrate the utility of functional near-infrared spectroscopy (fNIRS) as a tool for physicians to study cortical plasticity in children with cerebral palsy (CP). Motor cortex activation patterns were studied in five healthy children and five children with CP (8.4±2.3 years old in both groups) performing a finger-tapping protocol. Spatial (distance from center and area difference) and temporal (duration and time-to-peak) image metrics are proposed as potential biomarkers for differentiating abnormal cortical activation in children with CP from healthy pediatric controls. In addition, a similarity image-analysis concept is presented that unveils areas that have similar activation patterns as that of the maximum activation area, but are not discernible by visual inspection of standard activation images. Metrics derived from the images presenting areas of similarity are shown to be sensitive identifiers of abnormal activation patterns in children with CP. Importantly, the proposed similarity concept and related metrics may be applicable to other studies for the identification of cortical activation patterns by fNIRS. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3412746]

Keywords: functional near-infrared spectroscopy; cerebral palsy; plasticity; motor cortex; similarity; neuroimaging.

Paper 09553R received Dec. 14, 2009; revised manuscript received Mar. 9, 2010; accepted for publication Mar. 23, 2010; published online May 26, 2010.
tive manner as well as assess the response to a specific treatment.

Functional near-infrared spectroscopy (fNIRS) offers great promise in becoming the imaging tool of choice for physicians to monitor cortical activation patterns in children with CP. It is cost effective, safe, noninvasive, and less affected by movement than other imaging modalities, such as fMRI.15 fNIRS detects changes in light absorption and scattering in tissue caused by the changes in concentration of oxyhemoglobin (HbO) and deoxyhemoglobin (Hb) secondary to neuronal activity, known as neurovascular coupling.16 Moreover, fNIRS has much higher temporal resolution, up to tens of milliseconds,17 compared to fMRI, which has a temporal resolution measured in seconds.18,19

In this work, we demonstrate how fNIRS image analysis of motor cortex activation patterns due to finger tapping enables identification of spatial and temporal metrics that can be used to differentiate between normal pediatric subjects and children with CP with high statistical significance. The definition and use of these metrics, namely the distance from center, area difference, duration, and time-to-peak, is described. In addition, a similarity concept is defined that enables identification of both areas that have similar activation patterns as that of the maximum activation area, as well as dissimilar activation patterns within that area, which are not readily discernible in unprocessed activation images. It is shown that some of the metrics quantified from similarity images can serve as sensitive discriminators of abnormal activation patterns in children with CP.

2 Materials and Methods

2.1 Subjects

Five pediatric controls, Subjects 1–5 (two female and three male, 8.4 ± 2.3 years old), and five children with hemiparetic CP. Subjects 6–10 (three female and two male, 8.4 ± 2.3 years old), were included in this study. All the controls were right handed. Only children with CP that had a successful anatomical magnetic resonance imaging scan performed, indicating a single subcortical lesion in one of the two brain hemispheres, were included in this study. Three of the children with CP, Subjects 6–8, had a subcortical lesion in the left hemisphere (LH), which affected movement of their right hand. Subjects 9 and 10 had a subcortical lesion in the right hemisphere (RH), which affected movement of their left hand. Informed consent was obtained from all subjects and their legal guardians. Except for Subjects 4 and 6, it was possible to attain a second visit for fNIRS measurements within two months of the first. These studies were performed under the approval of The University of Texas Southwestern Medical Center at Dallas (UTSW) Institutional Review Board protocol (IRB No.:042007-064).

2.2 Measurements

A continuous-wave (CW) fNIRS brain imager (CW-5, Techen Inc., Milford, Massachusetts) was used to map the HbO and Hb changes induced by motor cortex stimulation. The source-detector geometry is shown in Fig. 1(a). Eight detectors were placed over each brain hemisphere to cover a relatively large area of the motor cortex. The geometrical arrangement of sources and detectors was attached onto the subjects’ heads by perforated Velcro straps. Eight laser sources emitted at 690 nm and eight at 830 nm, such that each optical fiber bundle delivered light of both wavelengths at each source location [circles in Fig. 1(a)], simultaneously. Each source bundle had up to four detectors [squares in Fig. 1(a)] within a 3-cm distance, and each detector received signals from up to two source bundles. As a result, there were 28 possible source-detector channel combinations for each wavelength, as indicated by the straight lines connecting sources to detectors in Fig. 1(a). Activation in cortical areas within the probes’ field of view (6 × 20 cm) was monitored simultaneously by all source-detector pairs, because the CW-5 enables all laser sources to be on at the same time with distinct modulation frequencies (6.4–12.6 kHz, with an increment of 200 Hz). The backreflected light detected was sampled at a rate of
100.16 Hz, which was later demodulated and downsampled to 20.03 Hz to reduce data-set sizes.

In addition to measuring the cortical hemodynamic response by optical means, the finger-tapping, respiration, and cardiac pulsation patterns were also measured simultaneously for each subject. The finger-tapping rate was measured using a custom-built, capacitance-based tapping board that produced a bilevel voltage output for each tap. The respiration and cardiac pulsation measurements were amplified with a Brownlee Model 410 amplifier (AutoMate Scientific, Berkeley, California), and recorded by the CW-5 system. The respiration and cardiac measurements were amplified by a factor of 200 and 30 mV/mV, respectively. Breathing patterns were measured using a respiration belt with a piezoelectric transducer (Sleepmate Technologies, Glen Burnie, Maryland). The belt was wrapped around each subject’s lower chest while keeping the transducer away from the heart in order to prevent cardiac pulsation noise from leaking into the respiration measurement. Cardiac pulsation was measured using a pulse oximeter (Nellcor Inc., Boulder, Colorado) attached to the index finger on the hand not in use for tapping. In addition, movement of the forearm flexor muscles controlling the fingers was measured by electromyography ([EMG], B & L Engineering, Santa Ana, California]. The EMG measurements were not used in our filtering procedures but instead served as quality control sentinel s for our measurements, making sure that the specified hand was the only hand tapping, while the other was kept still. A diagram of the overall instrumentation setup is shown in Fig. 1(b), in which EMG (R) and EMG (L) stand for the EMG taken from the right forearm and left forearm, respectively. All nonoptical physiological measurements were fed into the CW-5 and digitized by an analog-to-digital converter, sampled at 100.16 Hz and later downsampled to 20.03 Hz, to reduce data-set sizes, as was done for the reflectance signals. Therefore, all signals had a common time base with the fNIRS measurements, which enabled the use of the reference physiological measurements in the subsequent filtering procedures described in Section 2.5.

2.3 Protocol

The fNIRS probe placement was done according to the measured coronal (ear-to-ear) and sagittal (forehead-to-back-of-head) distances. The center of the probe set was placed at half the distance of the aforementioned measurements, which was considered the estimated midpoint of the motor cortex. A PowerPoint animation on a computer screen guided subjects through the finger-tapping protocol. No set tapping frequency was required of the pediatric subjects. Thus, the tapping frequency was self-paced, which resulted in average finger-tapping rates in the 1–2 Hz range within the 15 s tapping period. Tapping consisted of moving all fingers of one hand, while the wrist was held down by soft straps to keep it from moving. The subjects touched the tapping board while tapping and rested their hand on the board during the no-tapping intervals. The data acquisition protocol consisted of a 30-s baseline (no tapping), immediately followed by a series of ten consecutive epochs of 15 s of tapping and 25 s of rest, and ended with a 20-s baseline measurement, resulting in a total acquisition time of 450 s. Measurements were separately acquired for both left and right finger tapping. The subjects sat up straight with their head resting back in a quiet, dimly lit room. All subjects were video recorded during the measurements in order to verify, along with the aforementioned EMG measurements, that they performed the finger-tapping tasks properly.

2.4 Visualization of fNIRS Signals

Processing and visualization of the time-series reflectance data acquired by the CW-5 system is usually performed by the open-source HOMER software implemented in MATLAB (Mathworks, Natick, Massachusetts).20 HOMER filters the signal using low-pass and high-pass Butterworth filters to reduce noise from frequencies outside the known physiological range. Furthermore, the software applies principle component analysis (PCA) to reduce any global hemodynamic signals, detrends signals to reduce any long-term drifts, and averages the detected signal for individual source-detector pairs over user-selected time intervals of tapping and rest. Subsequently, using the reflectance measurements from all detector channels activation images are reconstructed by use of the Tikhonov perturbation solution to the photon-diffusion equation,21 which employs a localized Moore–Penrose inversion scheme.22 The reconstructed, two-dimensional images (21 × 20 pixel) represent maps of Hb and HbO changes on the brain cortex surface, within the detector’s field of view [rectangle in Fig. 1(a)] as a result of finger tapping. As Hb dynamics are highly correlated with those of HbO23 and have significantly lower amplitudes than the latter, making them susceptible to cross-talk from HbO24 and to interference from physiological artifacts, we have focused on the analysis of HbO dynamics only.

2.5 Filtering

In addition to detecting evoked hemodynamic changes, fNIRS is sensitive to cerebral hemodynamic fluctuations of systemic origin. Such systemic fluctuations can be caused by cardiac pulsation, respiration, and Mayer waves. For typical motor cortex activation protocols, the cortical hemodynamic response can be found in the 0.1–0.4 Hz frequency range, while physiological artifacts, such as cardiac pulsation, can be found between 0.8–2.0 Hz, respiration in the 0.1–0.3 Hz range, and Mayer waves at ~0.1 Hz or lower.25,26 Because there is a significant overlap between the frequency spectra of respiration and Mayer waves and of the hemodynamic response due to brain activity, and there is an overlap in frequency spectra of cardiac pulsation and the tapping frequency, bandpass filtering is not effective in removing such physiological artifacts. Previous attempts in the removal of these physiological artifacts from fNIRS signals have used component analysis27–30 or adaptive filtering.30–34 Two types of component analysis have previously been used to filter fNIRS signals, PCA,27 and independent component analysis.28–30 Component analysis is advantageous because no additional measurements are used to track the physiological artifacts. The disadvantage of these algorithms is that they need training data to properly identify the physiological artifacts and that the physiological artifacts may contribute to the signal differently once the study begins, thus making it difficult to attribute each component to a specific artifact or to define the number of components needed to be removed.27
To better track the physiological artifacts, additional physiological measurements can be simultaneously taken or the physiological artifacts and the hemodynamic response can be modeled, as in adaptive filtering. The least mean square (LMS) filter is a finite impulse response (FIR) filter in which the filter coefficients adjust in order to minimize the least-squared error in fitting the measured noise measurements to the noisy signal. The extended Kalman filter, otherwise, can filter the fNIRS signals by statistically modeling the physiological artifacts and the hemodynamic response due to brain activity. Unfortunately for the extended Kalman filter, a model of the hemodynamic response for children with CP is unknown and may vary, depending on the type of lesion the child has.

In order to overcome the disadvantages of individually filtering with component analysis or adaptive filtering, a combination of PCA and LMS filtering was used in this work. Using the additional reference physiological measurements, physiological artifacts that were within the same frequency range as the tapping frequency or the cortical hemodynamic response were filtered. More specifically, the fNIRS signals were first bandpass filtered between 0.01 and 2 Hz, using a Butterworth filter, to exclude frequencies of nonphysiological origin from the signal. Subsequently, baseline correction compensated for the baseline drift that could occur throughout the repeated cycles of activation and then PCA was applied to the fNIRS signals. In contrast to the baseline procedure implemented in HOMER, which uses only the first time point of the experiment as baseline reference, we corrected the baseline for each activation interval individually by taking the average of the 5 s of data before the beginning of each activation interval and then subtracting that value from the activation data in that interval. In the implementation of PCA, we empirically found that removing the two largest eigenvalues, which represented components in the frequency range of Mayer waves and cardiac pulsation physiological artifacts, resulted in the best compromise between the reducing global background signal while not substantially reducing the contrast of activation areas in the reconstructed images as was done by Zhang et al. In our analysis, we saw that the percent of the first two eigenvalues to the sum of all eigenvalues represented most of the baseline signal (93.04% ± 3.88%). In addition, in some cases the second eigenvector contributed to larger variations in the baseline signal than the first eigenvector. It was thus decided that the removal of the first two eigenvectors was optimal for removing baseline noise without compromising signal quality. Finally, the signal was adaptively filtered for respiration first, and then cardiac pulsation, using an LMS algorithm. The reference noise signals that were used in this adaptive filtering procedure were first low-pass filtered at 5 Hz to reduce the contribution of electronic noise. No difference was found in the order in which the adaptive filters were applied, but using PCA after adaptive filtering was found to greatly decrease the signal quality.

2.6 Averaging

After filtering the signals as described in Section 2.5 and excluding the tapping-rest intervals that were not executed properly, as judged by video recording and EMG, data from all source-detector pairs were processed by HOMER. Because the sampling frequency was 20.03 Hz, each time point represents every 0.05 s. The properly executed 40-s tapping-rest intervals were averaged at each time point (total of 800 time points) giving a 40-s averaged temporal response for each source-detector pair. Activation images of the motor cortex were then reconstructed for every time point from the averaged temporal response of all source-detector pairs. Time-averaged HbO activation images were formed by averaging HbO activation images within a window that started 5 s after the tapping began to 5 s after the tapping had stopped (5–20 s relative to the beginning of each 15-s tapping interval) because that was the time window during which most of the activation signal was observed.

2.7 Identifying Areas of Activation

Activation areas were identified by a k-means clustering algorithm implemented in MATLAB. The algorithm was applied to identify areas of activation in two different images: the time-averaged HbO images and the two-dimensional spatiotemporal plots representing a concatenation of the 40-s averaged temporal responses of each pixel.

The algorithm employed three cluster means that were initialized at the maximum pixel value, minimum pixel value, and zero. The algorithm then placed each image pixel into one of the clusters, as judged by the closest Euclidean distance [Eq. (1)] of the value of each pixel (of coordinates $x, y$) to the mean of the three clusters. However, if the minimum distance was equal between two clusters, then the pixel was placed into the cluster with the smaller mean.

$$\text{Distance}_{\text{cluster}}(x,y) = |\text{pixel\_value}(x,y) - \text{mean}_{\text{cluster}}|. \quad (1)$$

The algorithm recalculated the mean of each cluster after processing the whole image and repeated the process until the means of each cluster no longer changed, in which a maximum of 11 iterations were necessary for the time-averaged HbO images. The cluster that initially had a mean equal to the maximum pixel value in the image represented pixel values related to activation [Fig. 2(a)]. The cluster that initially had a mean equal to zero represented pixel values related to the baseline. Lastly, the cluster that initially had a mean equal to the minimum pixel value in the image represented deactivation.

Similarly, the temporal responses of each pixel were clustered to identify activation, baseline, and deactivation. The 40-s averaged temporal responses of each pixel were concatenated from the LH to RH to give a two-dimensional plot, in which the y-axis represented the pixels and the x-axis represented time. Then the previously discussed k-means clustering algorithm was applied to the two-dimensional plot, in which a maximum of 20 iterations were needed to identify the three clusters. The pixels with activation in the two-dimensional plot [Fig. 2(b)] showed the time and location of areas of activation during finger tapping. Subsequently, a number of spatial and temporal metrics were derived from the resulting time-averaged and two-dimensional activation images, as described here below.

2.8 Metrics Derived from the Reconstructed Images

We identified spatial metrics (distance from center and area difference) from the time-averaged HbO images and temporal
metrics (duration and time-to-peak) from the two-dimensional plots. The identified spatial and temporal metrics presented differences between controls and children with CP. We have also identified similar temporal patterns spatially over the motor cortex, as will be shown in Section 3, which has great promise for differentiating between the two subject groups.

2.8.1 Distance from center

To show differences in activation location in the time-averaged HbO images between healthy controls and children with CP, a vertical line was used as a reference at the center of the motor cortex and the horizontal distance between the near edge of the activation area closest to the centerline (distance from center) was measured for each hemisphere [Fig. 2(a)]. Since the distance from center metric is dependent on the identification of areas of activation, this metric was not quantified in the cases where there was no activation identified in the ipsilateral or contralateral hemisphere relative to the tapping hand.

2.8.2 Area difference

Because it was observed that children with CP consistently had activation closer to the motor cortex centerline, a metric was defined to quantify the amount of activation observed by this difference. The metric was defined as the difference in area of activation (area difference) occurring near the edge of the contralateral hemisphere of the tapping hand and the middle region of the motor cortex. More specifically, the detector field of view was divided in three nearly equal areas of 6 cm in length and a width of 7 cm for the RH and LH and 6 cm for the motor cortex center [Fig. 2(c)]. Areas of activation in each subdivision of the motor cortex were identified using the same clustering algorithm previously used to identify activation over the entire motor cortex (see Section 2.7). The metric of interest for differentiating between the two subject groups was defined as the difference of the identified activation area found by the clustering algorithm in the time-averaged HbO images between the contralateral hemisphere and the middle region of the motor cortex.

2.8.3 Duration and time-to-peak

The temporal aspects of the change in HbO were quantified with metrics (duration and time-to-peak) devised from the two-dimensional plots [Fig. 2(b)]. Lines of activation in these two-dimensional plots represented the change of HbO in each pixel that was above baseline, as identified by the activation cluster from the clustering algorithm. The length of these lines therefore represented the duration of activation for each pixel [Fig. 2(b)]. The metric of duration was then defined as the average value of activation duration over all pixels. Addition-
ally, the 40-s average temporal response data were sorted to identify pixels with maximum activation from the two-dimensional plots, as identified by the clustering algorithm. From the pixels with maximum activation, the average time difference between the beginning of the tapping interval and the time point of maximum HbO change was defined as the time-to-peak metric [Fig. 2(d)].

2.8.4 Similarity

Image areas, which were spatially distinct from the primary activation region, were observed to have demonstrated temporal profiles of HbO change similar to those of the primary activation region. Because the amplitudes in HbO change of these secondary regions were lower in magnitude, they were not always assigned to the cluster of pixels with activation, as identified by the clustering algorithm discussed above.36 To identify these areas, a similarity concept was defined that quantified aspects of resemblance between the temporal profile of the maximum activation area and other regions in the motor cortex.

Only pixels that had a value greater than or equal to zero from the time-averaged HbO activation images were compared in the similarity analysis. Before checking the similarity between the temporal profiles of each pixel, these were first smoothed to remove high-frequency instrumentation noise using a third-order Savitzky–Golay FIR filter.38,39 After smoothing, each pixel was normalized to the peak activation value during the 40-s tapping-rest interval because the similarity analysis focuses on the shape of activation curves and not their magnitude.

Subsequently, two metrics were defined for describing each pixel’s temporal activation profile. These were the angle defined by the tangent (dθ) and the change of signal amplitude (ds) at every time point, as shown in Fig. 3(a).40 A cost function was then defined to quantify the degree of similarity for these two metrics between the seed pixel (i.e., the peak activation pixel) and other pixels in the image. Though previous works have used the Euclidean distance to find the similarity between two time series,41–43 a dynamic time-warping (DTW)–based cost function was selected to measure similarity to accommodate for small changes in phase between the seed pixel and activation occurring elsewhere in the image.44,45 A brief overview of how DTW was adapted to the fNIRS image analysis is explained next.

Suppose there are two time series, S and P, of length N and M, respectively, where $S=s_1, s_2, \ldots, s_N$ and $P=p_1, p_2, \ldots, p_M$. An $N \times M$ matrix is then formed to align the two sequences in which the entry $(n,m)$ consists of the distance, $d(s_n, p_m)$, of the two points $s_n$ and $p_m$ [i.e., $(s_n - p_m)^2$]. A warping path45 is then found in which a contiguous set of matrix elements defines a mapping between S and P [Fig. 3(b)] and that the warping cost, $DTW(S, P)$, is minimized such that

$$DTW(S, P) = \min \left[ \sqrt{\sum_{k=1}^{K} w_k} \right],$$

where $K$ is the number of warping elements and $w_k$ is the distance found at $(n,m)$. The warping path is subject to specific constraints such that the warping starts and finishes in diagonally opposite corners, steps are only allowed to adjacent cells (including diagonally adjacent cells), and each $w_k$ must be monotonically spaced in time.45 Furthermore, the time warping is globally constrained using the Sakoe–Chiba band,46 where $r=10$ (i.e., 0.5 s). The purpose of this constraint is to prevent a relatively small section of one time series to be mapped to a relatively large section of another. In this case, a maximum of 0.5 s of one time series can be mapped to a single point of another (i.e., a shift of 0.5 s can be detected).

In this study, DTW was not directly applied to the pixel temporal responses, but to the two metrics that defined the shape of the temporal responses, $d\theta$ and $ds$. $N \times M$ matrices were formed to find the warping costs for both $d\theta$ and $ds$. Instead of looking for the minimum warping cost for $d\theta$ and $ds$ separately, the warping cost was redefined such that $w_k$ in Eq. (2) is represented by

---

Journal of Biomedical Optics 036008-6 May/June 2010 • Vol. 15(3)

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics on 07 Jun 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]

Subsequently, a threshold needed to be defined to identify the range of DTW cost corresponding to activation image pixels that were considered similar to the seed pixel. Because the dynamic warping cost is directly comparable between controls and subjects with CP, the dynamic warping costs of the controls, for both right and left finger tapping, were pooled together in a single histogram that presented the frequency of each dynamic warping cost for all control subjects [e.g., Fig. 3(c)]. Because there were no clear divisions found in the histogram, a $k$-means algorithm was used to divide the histogram into five levels, in which the clusters were initialized similar to that of Liu and Yu. Though the $k$-means algorithm could divide the histogram into more levels for a more stringent similarity definition, dividing it into five levels was empirically found to be sufficient for finding similar time series, as will be shown in Section 3. A flowchart of the similarity algorithm is shown in Fig. 4.

\[ w_k = (d(\theta_n - \theta_{n+1})_k^2 + (d(s_n - s_{n+1})_k^2. \quad (3) \]
ipsilateral hemisphere than in the contralateral hemisphere relative to the tapping hand. In contrast, the ipsilateral activation for children with CP was closer to the centerline, and in a few cases, it was even closer than the contralateral activation for the same subject.

A high statistical significance was found ($p=0.0008$, $1-\beta=0.98$, Table 1) for the mean value difference in ipsilateral distance from center between controls (7.29 cm ± 1.98 cm) and children with CP (3.14 cm ± 1.46 cm), but no significant difference ($p=0.54$) was observed in the contralateral distance from center between controls (3.80 cm ± 1.48 cm) and ones with CP (3.30 cm ± 2.06 cm).

### 3.2 Area Difference

Although the distance from center metric gave information about differences in activation location between controls and children with CP, this metric did not capture differences in the spatial extent of the corresponding activation areas. To quantify the latter, the area-difference metric was defined as described in Section 2.8.2.

In Fig. 6, the results of the area difference metric show that for controls the area of activation in the contralateral hemisphere of tapping was always greater than the activation in the middle area of the motor cortex by at least 2 cm$^2$. Three of the five subjects with CP (Subjects 6, 9, and 10) showed greater activation area in the middle portion of the motor cortex when tapping with one of their two hands. Interestingly, Subject 8, a right hemiparetic subject, presented an area difference of <1 cm$^2$ when right finger tapping (i.e., even though he was using his affected hand). Subject 7 had similar results as the controls for this metric. Although in some cases, subjects with CP presented larger activation in the middle region of the motor cortex than in the contralateral hemisphere, for the current population size there was no significant

![Fig. 6](https://example.com/fig6.png) Activation area difference for control subjects (Subj1–Subj5, left tapping–black, right tapping–white), ones with right hemiparetic CP (Subj6–Subj8, left tapping–brick, right tapping–dotted), and ones with left hemiparetic CP (Subj9–Subj10, left tapping–trellis pattern, right tapping–horizontal lines).

![Fig. 7](https://example.com/fig7.png) Bar graph presenting the ratio of duration over time-to-peak metrics for control subjects (Subj1–Subj5, left tapping–black, right tapping–white), ones with right hemiparetic CP (Subj6–Subj8, left tapping–brick, right tapping–dotted), and ones with left hemiparetic CP (Subj9–Subj10, left tapping–trellis pattern, right tapping–horizontal lines).

### Table 1

<table>
<thead>
<tr>
<th>Control RFT/LFT versus affected</th>
<th>Distance from center</th>
<th>Area difference</th>
<th>Duration/time-to-peak</th>
<th>Area difference (similarity)</th>
<th>Distance from seed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity (%)</td>
<td>100</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>80</td>
<td>60</td>
<td>100</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>100</td>
<td>60</td>
<td>83</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>83</td>
<td>60</td>
<td>100</td>
<td>66</td>
<td>71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control RFT/LFT versus unaffected</th>
<th>Distance from center</th>
<th>Area difference</th>
<th>Duration/time-to-peak</th>
<th>Area difference (similarity)</th>
<th>Distance from seed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity (%)</td>
<td>100</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>60</td>
<td>20</td>
<td>60</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>100</td>
<td>33</td>
<td>75</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>71</td>
<td>43</td>
<td>66</td>
<td>50</td>
<td>71</td>
</tr>
</tbody>
</table>
difference ($p=0.068$) in area difference metrics between controls ($9.04 \text{ cm}^2 \pm 5.96 \text{ cm}^2$) and children with CP ($3.40 \text{ cm}^2 \pm 7.01 \text{ cm}^2$). Also, there was no statistically significant difference in the mean values of the area difference metric when subjects with CP were tapping with their affected ($0.53 \text{ cm}^2 \pm 5.81 \text{ cm}^2$) or unaffected hand ($6.27 \text{ cm}^2 \pm 7.51 \text{ cm}^2$, $p=0.2135$).

### 3.3 Duration and Time-to-Peak

The ratio of duration over time-to-peak was identified as a potentially useful metric for differentiating between the two subject groups. With one exception, the controls had a duration over time-to-peak ratio of $\geq 1$ for both right and left finger tapping (Fig. 7), i.e., duration was greater than or equal
to time-to-peak. In contrast, duration was found to be less than the time-to-peak when children with CP were tapping with their affected hand (right, dotted column for Subjects 6–8 and left, trellis patterned column for Subjects 9 and 10). In addition, when children with CP were tapping with their unaffected hand, this ratio often assumed values that were substantially higher or lower than unity for some individuals, though no consistent pattern was observed. No significant difference ($p=0.99$) was observed between the mean values of duration over time-to-peak ratios between controls that were left finger tapping ($1.19 \pm 0.09$) or right finger tapping ($1.19 \pm 0.26$). Subsequently, these results were pooled into a single normal subject group and a two-sampled t-test verified the high statistical significance ($p=0.0002$, $1-\beta=0.99$, Table 1) in the difference between the means of the duration over time-to-peak ratios between controls ($1.19 \pm 0.18$) and children with CP tapping with their affected hand ($0.60 \pm 0.25$).

### 3.4 Similarity

By using the above-described similarity algorithm (Section 2.8.4), pixel areas were identified having similar temporal profiles of HbO concentration change with respect to the primary area of activation. It was found that these similar areas did not necessarily coincide with the areas having the highest activation in the time-averaged images, but could have an HbO concentration change in the baseline range. Furthermore, subjects with CP demonstrated a greater diversity of temporal patterns, as quantified by the range of DTW penalty term values, within the areas of activation than control subjects. To illustrate these points, a sample activation image [Fig. 8(a)], an image processed by an activation cluster algorithm [Fig. 8(b)], an image processed by the similarity algorithm [Fig. 8(c)], and the color-coordinated normalized, average temporal profiles of the areas in the similarity image [Fig. 8(d)] are shown for a control subject. For purposes of comparison, the same types of images are shown for a subject with CP [Figs. 8(e)–8(h)]. It is seen that the area of activation identified by the clustering algorithm [Fig. 8(b)] does not coincide with the area of pixels with similar activation [Fig. 8(c)]. In contrast, children with CP had areas of activation [Fig. 8(e)] with dissimilar activation profiles, except for pixels immediately neighboring the seed-pixel location [Fig. 8(g)]. The temporal profiles of activation are also shown [Figs. 8(d) and 8(h)], in order to provide a visual example of the degree of similarity between the seed pixel (black curves) and the most similar cluster of pixels (level 1). Because the dynamic warping costs of the control subjects were pooled together, not all subjects individually present all five levels of similarity as seen in Figs. 8(c) and 8(d).

The area difference of the most similar pixels to the primary activation area (level 1) in the similarity images was investigated as a metric for differentiating between controls and children with CP. In Fig. 9, it is seen that controls typically had a greater area difference than children with CP. This difference between the mean values of area difference of controls ($12.14 \text{ cm}^2 \pm 12.85 \text{ cm}^2$) and children with CP ($0.29 \text{ cm}^2 \pm 4.66 \text{ cm}^2$) was indeed found to be significant ($p=0.0383$, $1-\beta=0.77$, Table 1). The results for area difference demonstrate that the similarity concept can highlight differences between controls and subjects with CP much better than the corresponding metric derived from applying a threshold algorithm to the same subject’s activation images. Additionally, there was no discernible correlation between area difference and whether children with CP were tapping with their affected hand ($1.82 \text{ cm}^2 \pm 5.09 \text{ cm}^2$) or unaffected hand ($1.25 \text{ cm}^2 \pm 4.14 \text{ cm}^2$, $p=0.3259$).

Although children with CP had smaller areas of similarity, their most similar areas (level 1) were spatially closer to their seed pixel than in controls (Fig. 10). To present this difference the maximum distance among the most similar areas (level 1) from the seed pixel was found, in which the difference between the controls ($7.43 \text{ cm} \pm 6.00 \text{ cm}$) and children with CP ($2.18 \text{ cm} \pm 1.17 \text{ cm}$) was significant ($p=0.0223$, $1-\beta=0.73$, Table 1). The larger distance in controls can be explained by some similarity areas being seen in the ipsilateral hemisphere and the larger areas of similarity neighboring the seed pixel, whereas in children with CP this is not the case. Furthermore, there is no significant difference ($p=0.7427$) when the children with CP were tapping with their...
affected (2.05 cm ± 1.53 cm) or unaffected hand (2.31 cm ± 0.84 cm).

3.5 Validation Analysis to Differentiate between Controls and Subjects with CP

Though four of the five presented metrics gave statistically significant differences between controls and children with CP (distance from center, duration/time-to-peak, area difference of similarity images, and maximum distance from seed), there were still false positives or false negatives found. Table 1 shows the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each of these metrics. Subjects were considered to be in the CP group if their metrics were beyond one standard deviation of the corresponding mean value of the same metric for controls.

From Table 1, it is clear that each fNIRS metric performs differently in each category. For example, the specificity, sensitivity, PPV, and NPV values for area difference in activation images are low, which supports the t-test conclusion (p-value = 0.068) that this metric may be less useful for differentiating between the two subject populations. On the other hand, distance from center and distance from seed clearly tell if a subject is a control (PPV = 80%, NPV = 100%), while duration/time-to-peak clearly tells if a subject is in the CP group (PPV = 100%, NPV = 80%). However, the area difference of similarity does not clearly show if a subject is a control or with CP, but it does have a higher specificity (80%) than the area difference in activation images (60%). Because no single metric has 100% success, an approach for distinguishing between the two subject populations where multiple metrics are quantified is warranted.

| Table 2 | The ipsilateral distance from center and area difference metrics measured in two separate visits for four subjects in each group. |

<table>
<thead>
<tr>
<th>Control</th>
<th>Ipsilateral distance from center (cm)</th>
<th>Area difference (cm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Visit 1</td>
<td>Visit 2</td>
</tr>
<tr>
<td>LFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Subject 2</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Subject 3</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Subject 5</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>RFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 1</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Subject 2</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Subject 3</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Subject 5</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>mean ± std.</td>
<td>(7.83 ± 1.47)</td>
<td>(7.00 ± 1.10)</td>
</tr>
<tr>
<td>Paired T test</td>
<td>p = 0.2940</td>
<td>p = 0.3181</td>
</tr>
<tr>
<td>LFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 7 (RH)</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>Subject 8 (RH)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Subject 9 (LH)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Subject 10 (LH)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>RFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 7 (RH)</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Subject 8 (RH)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Subject 9 (LH)</td>
<td>X</td>
<td>3</td>
</tr>
<tr>
<td>Subject 10 (LH)</td>
<td>X</td>
<td>5</td>
</tr>
<tr>
<td>mean ± std.</td>
<td>(2.80 ± 1.30)</td>
<td>(3.62 ± 1.69)</td>
</tr>
<tr>
<td>Paired T test</td>
<td>p = 0.8350</td>
<td>p = 0.7231</td>
</tr>
</tbody>
</table>

X is shown when no activation was found in the ipsilateral hemisphere.
3.6 Repeatability

The exact anatomical placement of probes and the quality of optical contact with the scalp are common causes of concern for the repeatability of fNIRS measurements. To reduce such concerns in this study, measurements were performed twice for four subjects in each group (Table 2). As an example, the reflectance time series for the source-detector pair with maximum signal, which was the same for both visits, are presented in Figs 11(a) and 11(b) for a control subject (Subject 2) and for a subject with CP (Subject 8) tapping with their right hand. It is seen that these time-series signals, as well as the reconstructed activation and activation-threshold images [Figs. 11(c) and 11(d)] have similar patterns for both visits.

Furthermore, spatial metrics (distance from center and area difference) were quantified for both visits and a paired t-test was performed on these metrics for each subject group. For both groups, there was no significant difference in these metrics between the two visits (Table 2, X is shown when no activation was found in the ipsilateral hemisphere). The repeatability findings were similar for all other metrics (not shown).

4 Discussion and Conclusions

A CW fNIRS instrument was used to image patterns of change in HbO concentration in the motor cortex of five con-
trol children and five children with CP for a finger-tapping protocol. A signal-processing method, involving a combination of bandpass filtering, PCA, and adaptive filtering, was developed to remove global physiological artifacts, such as respiration, cardiac pulsation, and Mayer waves. The 40-s tapping-rest intervals were averaged together at each time point giving a 40-s averaged tapping-rest response for each source-detector pair. Images were subsequently reconstructed for each time point (0.05 s) from the averaged temporal responses of all source-detector pairs. These series of images, or their corresponding time-averaged images, were processed by a clustering algorithm to identify the location and spatial extent of areas of activation. Additionally, the same reconstructed images were analyzed by a similarity algorithm that identified areas of lower change in HbO concentration compared to primary activation areas that nevertheless had similar temporal change patterns to the latter. Therefore, the similarity algorithm unveiled secondary activation areas that to our knowledge have not been studied before. Because of their smaller activation values, compared to the primary activation region, these areas were not easily discernible in the time-averaged activation images and were considered to be in the range of the baseline data according to the clustering algorithm.

After application of the activation threshold or of the similarity algorithm, a number of metrics were quantified in the processed images. These metrics were either spatial (distance from center and area difference) or temporal (duration and time-to-peak) in nature. The thresholded activation images presented differences between controls and children with CP using the distance from center metric and the ratio of duration over time-to-peak. Similarity analysis provided the area difference metric (Fig. 9) with higher statistical significance for differentiating between controls and children with CP compared to the corresponding metric from thresholded time-averaged images of activation (Fig. 6). Moreover, similarity displayed a significant difference in spatial location of areas of similarity by measuring the maximum distance of the most similar areas of the seed pixel from the seed pixel. It was found that some of the above-described metrics could differentiate between controls and children with CP with high statistical significance and statistical power, despite the small numbers of subjects involved, as summarized in Table 3. It is important to point out that the identified differences between subject populations for any one metric were deduced by statistical analysis. Therefore, a subject with CP may have had some metrics with values that overlapped with the range of corresponding values for control subjects. It is conceivable that as the subject population increases with future fNIRS studies, it will be possible to differentiate more clearly between subject groups by training a classification algorithm (e.g., a support vector machine or a neural network) to accept as an input a set of fNIRS image metrics for each individual and correctly classify them.

Although fNIRS may not be needed for the initial diagnosis of the CP condition, it can become a practically implementable method that will enable clinicians to study cortical plasticity. The analysis methods presented above offer quantitative tools that may help differentiate between individuals that appear to have the same CP disability phenotype as quantified by standard upper extremity motor skill classifications schemes (e.g., MACS, SHUEE, and Nine Hole Peg Test). In addition, the current work is a first step toward the endeavor to establish fNIRS as an easily accessible tool for monitoring plastic changes in the cortex of patients with CP during and after therapy, which should be feasible as recent fMRI studies have indicated.

Acknowledgments
This work was supported by a grant from the William Randolph Hearst Foundation and The United Cerebral Palsy Research and Educational Foundation. We also acknowledge the clinical staff at the Texas Scottish Rite Hospital in Dallas for its time and support. We show our gratitude to Dr. George Kondraske of the University of Texas at Arlington for providing the tapping board.

References

Table 3 Summary of results of the differences found between controls and children with CP.

<table>
<thead>
<tr>
<th>Metric</th>
<th>p value</th>
<th>Statistical power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from center</td>
<td>&lt;0.0001</td>
<td>0.98</td>
</tr>
<tr>
<td>(Ipsilateral hemisphere of activation image)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration/time-to-peak</td>
<td>0.0002</td>
<td>0.99</td>
</tr>
<tr>
<td>(CP affected hand, activation image)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area differencea</td>
<td>0.0383</td>
<td>0.77</td>
</tr>
<tr>
<td>(Similarity image)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum distance from the seed pixela</td>
<td>0.0223</td>
<td>0.73</td>
</tr>
<tr>
<td>(Similarity image)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aIndicates the use of the Mann-Whitney test instead of a two-sampled test.