Dear Author,

YOUR PAGE PROOF IS AVAILABLE IN PDF FORMAT; please refer to this URL address http://115.111.50.156/jw/AuthorProofLogin.aspx?pwd=9c6541504534

Login: your e-mail address
Password: 9c6541504534

The site contains 1 file. You will need to have Adobe Acrobat Reader software to read these files. This is free software and is available for user downloading at http://www.adobe.com/products/acrobat/readstep.html.

This file contains:

Adobe Acrobat Users - NOTES tool sheet

A copy of your page proofs for your article

Please read the page proofs carefully and:

1) indicate changes or corrections using e-annotation;
2) answer all queries;
3) proofread any tables and equations carefully;
4) check that any special characters have translated correctly.

Special Notes:

Your prompt attention to and return of page proofs will help expedite publication of your work. Thank you for your cooperation.

Return to:

Mary Obrien, Journal Production Manager
Mary.Obrien@cenveo.com

If you have any questions regarding your article, please contact me. PLEASE ALWAYS INCLUDE YOUR ARTICLE NO. (MED-11-1039) WITH ALL CORRESPONDENCE.

This e-proof is to be used only for the purpose of returning corrections to the publisher.
Sincerely,

Mary Obrien, Journal Production Manager
E-mail: Mary.Obrien@cenveo.com
Dear Author,

Below are the queries associated with your article; please answer all of these queries before sending the proof back to Cenveo. Production and publication of your paper will continue after you return corrections or respond that there are no additional corrections.

<table>
<thead>
<tr>
<th>Location in article</th>
<th>Query / Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQ1</td>
<td>Sections 2.1-2.5 have been renumbered in order to numerically include the section “Data Collection and Analysis.” Please check edits of the same.</td>
</tr>
<tr>
<td>AQ2</td>
<td>Equations numbering has been changed to fit ASME standards, and captions have been removed. At times, the caption has been rephrased to fit into the paragraph above the equations. Please check edits of the same.</td>
</tr>
<tr>
<td>AQ3</td>
<td>Please provide DOI for the refs. 1, 11, 23, 26, 33, and 34.</td>
</tr>
</tbody>
</table>

Thank you for your assistance.
Assessment of Infant Movement
With a Compact Wireless Accelerometer System

There is emerging data that patterns of motor activity early in neonatal life can predict impairments in neuromotor development. However, current techniques to monitor infant movement mainly rely on observer scoring, a technique limited by skill, fatigue, and inter-rater reliability. Consequently, we tested the use of a lightweight, wireless, accelerometer system that measures movement and can be worn by premature babies without interfering with routine care. We hypothesized that this system would be useful in assessing motor activity, identifying abnormal movement, and in reducing the amount of video that a clinician would need to review for abnormal movements. Ten preterm infants in the NICU were monitored for 1 h using both the accelerometer system and video. A physical therapist trained to recognize cramped-synchronized general movements scored all of the video data by labeling each abnormal movement observed. The parameters of three different computer models were then optimized based on correlating features computed from accelerometer data and the observer’s annotations. The annotations were compared to the model’s prediction on unseen data. The trained observer identified cramped-synchronized general movements in 6 of the 10 infants. The computer models attained between 70% and 90% accuracy when predicting the same observer label for each data point. Our study suggests that mini-accelerometers may prove useful as a clinical tool assessing patterns of movement in preterm infants. [DOI: 10.1115/1.4006129]

Keywords: accelerometer, monitoring, infant movement, modeling

1 Introduction

The goal of this study was to begin to test the capability of minimally invasive, lightweight accelerometers to detect abnormal patterns of physical activity in prematurely born babies. Over the past two decades, the incidence and survival of preterm births (infants born at less than 37 weeks of gestation) have increased dramatically [1]. Not surprisingly, long-term neuromotor complications often associated with prematurity such as cerebral palsy (CP) [2,3] or the less severe category of minor neurological dysfunction [4] have increased as well. Prematurity is also associated with seemingly paradoxical detriments to body composition later in life, namely, obesity, on the one hand, and growth retardation with osteopenia on the other [5–7]. These abnormalities are tied to impaired energy balance (the relationship between food intake and energy expenditure such as physical activity). The developing brain, nerves, and muscle early in life are remarkably plastic [8–11], as are the metabolic mechanisms that control body composition [12]. Recent data
Additionally show that the quality of motor activity early in life in preterm infants is associated with school-age intelligence [13]. The early assessment of physical activity patterns in premature babies is increasingly recognized as an essential step in identifying metabolic and/or neuromotor impairments and optimizing therapeutic approaches [14–18].

Given the fragility of this population and the constraints imposed on diagnostic procedures in premature babies, it is not surprising that early assessment of physical activity has proven to be quite challenging. Such measures depend almost exclusively on direct observation of infants, or on viewing post hoc, real-time videotape of infant activity [19,20]. Very little is known about the normal developmental patterns of physical activity in premature babies, and, as a consequence, identification of abnormalities can occur quite late. For example, it is generally agreed upon that the diagnosis of a condition like CP cannot be made definitively until a child is at least 4 years-old [3]. While sophisticated brain imaging such as MRI can provide additional anatomic mechanisms for conditions like CP [21], these approaches are not yet suited for screening and early diagnosis.

There has been limited research using accelerometers to monitor infant movement. Factors such as accelerometer weight and size, which are not problematic for physical activity measurement in older children [22,23], are major obstacles in premature babies. Accelerometers have been used in infants to assess sleep/wake cycles [24], more accurately assess the physiology of swallowing [25], and identify movements that could confound assessment of oxygen saturation derived from cutaneous monitors [26]. In addition, several studies conducted by Pickler et al. [27,28] have been done to examine preterm infant movement and sucking. Very few studies, however, have quantifiably measured infant limb movements and tied these movements with neurodevelopmental outcomes. Intriguingly, even using a large (4 g, 20 × 12.5 × 7.5 mm) commercially available accelerometer placed on a single upper extremity, Ohgi et al. [29], in pioneering work, found different patterns of spontaneous movements of premature infants with known brain injuries compared with controls. In addition, promising research has recently been conducted by Heinze et al. [30] showing that a wired accelerometer could be used to differentiate between healthy babies and those at risk for cerebral palsy.

In order to optimize measurement and recording of spontaneous movements in premature babies, we recently developed an accelerometer small enough (13 × 10 × 8 mm) to be placed unobtrusively on all four extremities and the forehead of a premature infant and light enough (< 2 g) so that it would not interfere with spontaneous movement. In addition, the accelerometers are wireless, permitting real-time data acquisition and analysis [31]. Thus, we could measure the spontaneous movement of multiple extremities simultaneously.

As noted, a barrier in evaluating any new approach toward measuring infant movement is the dearth of metrics for comparison; however, in the case of the early diagnosis of CP, there exists a standardized, direct observation tool developed by Prechtl [32]. We designed the present experiments to compare our accelerometers with the Prechtl direct observational approach. In normal infants, Prechtl defined “general movements” (GMs) as elegant, smooth, variable in sequence, intensity and speed with a clear beginning and end. Prechtl also observed a unique abnormality of GMs that he named “cramped-synchronized” (CSGMs) in which the infant’s limbs were rigid and moved nearly in synchrony. CSGMs have high predictive value for the development of cerebral palsy [33-35]. We hypothesized that the accelerometers developed specifically for prematurely born infants could accurately identify CSGMs.

2 Methods

This protocol was reviewed and approved by the Human Subjects Institutional Review Board at UCI. To identify potential participants, we screened the medical records of infants in the NICU at the UCI Medical Center and recruited preterm infants with a gestational age at birth of between 23 and 36 weeks. Infants were excluded if they had mothers less than age 18 or if they had skin disorders, which could preclude the attachment of the accelerometers to the skin. The parents of ten premature infants provided written informed consent and enrolled in the study. Information regarding the participants is provided in Table 1.

All infants were monitored and videotaped for 1 h at 30-43 weeks corrected gestational age (see Table 1) in their isolette wearing only a diaper and with all swaddling removed to allow for free limb movement. The ambient temperature of the isolette was adjusted and maintained according to the judgment of the NICU nurse. A video camera was positioned with a mid-sagittal view of the infant above the isolette at a downward angle of 45° to record motion for post hoc video scoring.

Five accelerometers were used for data collection. Each one measured three orthogonal axes of acceleration on the head and each of the four limbs. Devices were embedded in cloth bands that were placed around the wrists, ankles, and forehead of the infants with a canonical anatomical orientation.

2.1 Data Collection and Analysis. The accelerometers transmitted data that was sampled nonuniformly at approximately 19 Hz in real-time to a computer located nearby. The raw accelerometer data consisted of real valued samples of the three axes measuring the degree of acceleration due to gravity and changes in limb motion. Although data was collected from four limbs and the forehead, subsequent analysis demonstrated that head movement did not substantially vary and was excluded from further analysis.

2.2 Data Annotation. The collected video data was annotated based on the visual observation technique developed by Prechtl et al. for identifying CSGMs. A NICU physical therapist, trained in the assessment, identified windows of time in the video in which the infants demonstrated CSGMs. The physical therapist annotated start and stop time for each CSGM observed. It is important to note that this physical therapist was employed at a different NICU, had no clinical contact with participants in the study and was masked to the infant’s medical history and age. This scoring technique provided a binary label for every time sample of either “CSGM Observed” or “CSGM Not Observed.” The annotated data was then synchronized with the corresponding timestamps in the video and accelerometer data.

2.3 Accelerometer Data Processing. The raw accelerometer data was collected at 19 Hz and was indexed by a time stamp. Each time stamp had associated with it 12 values corresponding to three axes of acceleration data collected for each of four limbs. This raw data required significant cleaning and processing to support effective modeling. The result of the data processing was to create 166 values associated with each time stamp. These were “features” that could be correlated with video observations at the

Table 1 Participant information. GA: gestational age

<table>
<thead>
<tr>
<th>ID#</th>
<th>GA at birth</th>
<th>Birth Wt.</th>
<th>GA at monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID 1</td>
<td>28</td>
<td>1480</td>
<td>30</td>
</tr>
<tr>
<td>ID 2</td>
<td>25</td>
<td>860</td>
<td>36</td>
</tr>
<tr>
<td>ID 3</td>
<td>36</td>
<td>2150</td>
<td>37</td>
</tr>
<tr>
<td>ID 4</td>
<td>24</td>
<td>880</td>
<td>34</td>
</tr>
<tr>
<td>ID 5</td>
<td>26</td>
<td>765</td>
<td>36</td>
</tr>
<tr>
<td>ID 6</td>
<td>25</td>
<td>600</td>
<td>43</td>
</tr>
<tr>
<td>ID 7</td>
<td>26</td>
<td>800</td>
<td>35</td>
</tr>
<tr>
<td>ID 8</td>
<td>25</td>
<td>680</td>
<td>39</td>
</tr>
<tr>
<td>ID 9</td>
<td>28</td>
<td>1150</td>
<td>36</td>
</tr>
<tr>
<td>ID 10</td>
<td>28</td>
<td>1180</td>
<td>37</td>
</tr>
</tbody>
</table>

AVERAGE 27.1 weeks 1063.5 gm 36.3 weeks
same time stamp and were the bases of the CSGM modeling (see Fig. 1).

There were three primary artifacts in our raw data that needed to be accounted for in our preprocessing. The first was that the data was not collected at uniform intervals. Although each reading was accurately time-stamped, the sampling frequency was irregularly spaced with a 19 Hz average. This made calculating features based on frequency analysis outside the scope of this study. The second was that every triple of accelerometer readings contained a 1-g component due to gravity, which was indistinguishable from a 1-g change in movement in a baby. Based on the orientation of the baby’s limbs this component was distributed across the three axes in different proportions at different times. The third was that there were periods of time in our collected data in which various NICU medical interventions invalidated a few seconds of our data. The impact of these artifacts was reduced as follows:

We manually identified periods of time in which the accelerometer data was not generated by spontaneous movements in the baby by reviewing the video. Examples of situations like this included pacifier adjustments by the NICU nurses, bumping of the isobed during temperature adjustments, and adjustments to the monitoring equipment. Removing this data resulted in gaps in the data stream that were conceptually similar to the already irregularly collected data samples although they were an order of magnitude larger. Each baby subsequently had a different absolute mean of a 10 s window centered on each data point. This is comparable to the removal of the DC component of the FFT used by Bao and Intille [36] given that regular samples were not available.

The low frequency filtering equations are as follows:

\[
 t_{ij} = t_i - m_{ij} \quad [i \in 1 : N, j \in x, y, z] \\
 m_{ij} = \text{mean}(t_{ij}) \quad [t - 5s : k : i + 5s] 
\] (1)

For every sample time \( i \) and every axis \( j \), we subtracted from the observed data point, \( t_{ij} \), the mean of a 10 s window centered at time \( i \). The number of data points over which the mean was calculated was not constant because the data was collected at nonuniform intervals and because some data was removed due to intervention. However, the mean was calculated over the same duration of time. This process served to eliminate the effect of gravity and any other systematic constant offset in our readings (e.g., calibration drift) without reducing the number of data points.

In this study, we did not model the posture of the baby, only the changes in the motion of the limbs, so we subsequently merged the three axes from each accelerometer into a single value that captured the overall magnitude of the smoothed data. For each time stamp, this resulted in four new data values, one that aggregated the motion from each limb. These were the first four features and represented the “corrected acceleration magnitude.”

\[
 F^2_i = \sqrt{(t_{i1})^2 + (t_{i2})^2 + (t_{i3})^2} \\
 L \in \{\text{left arm, right arm, left leg, right leg}\} 
\] (2)

One of the motivations for this method of data preprocessing (as opposed to aggregating over non-overlapping temporal windows) was to ensure that the number of data points was not reduced so that they could be matched to the video annotations with high temporal fidelity.

2.4 Feature Extraction. From the smoothed stream of accelerometer data, we next extracted statistical features for analysis. The features chosen were based on a desire to capture aspects of CSGMs specifically and general movements of the limbs more broadly.

We calculated the following instantaneous sample-based features for each data point:

- Maximum values of the corrected acceleration magnitude of the upper and lower body

\[
 F^5_i = \max(F^\text{left-arm}_i, F^\text{right-arm}_i) \\
 F^6_i = \max(F^\text{left-leg}_i, F^\text{right-leg}_i) 
\]

- Maximum of all the limbs

\[
 F^7_i = \max(F^\text{left-arm}_i, F^\text{right-arm}_i, F^\text{left-leg}_i, F^\text{right-leg}_i) 
\]

- Product of the corrected magnitude of the upper and lower body

\[
 F^8_i = F^\text{left-arm}_i \times F^\text{right-arm}_i \\
 F^9_i = F^\text{left-leg}_i \times F^\text{right-leg}_i 
\]

- Product of the corrected magnitude of the total body

\[
 F^{10}_i = F^\text{left-arm}_i \times F^\text{right-arm}_i \times F^\text{left-leg}_i \times F^\text{right-leg}_i 
\]

After these initial ten features, we also calculated several features based on functions that aggregated over windows of time: the mean, the maximum, the minimum, the standard deviation, and the z-value. The time windows we used covered 1 s, 2 s, or 4 s centered on the current data point. The data that we aggregated included each of the first ten features. This resulted in an additional 150 features (See Table 2).

The final six features that we calculated were the Pearson correlation between the left and right arm and the right and left leg. Although the correlation was for a single pair of data points, we used the same three windows as above to calculate the aggregate statistics required for each Pearson correlation calculation.

\[
 F^{161,162,163}_i = \text{Correl}_{1s,2s,4s}(F^\text{left-arm}_i, F^\text{right-arm}_i) \\
 F^{164,165,166}_i = \text{Correl}_{1s,2s,4s}(F^\text{left-leg}_i, F^\text{right-leg}_i) 
\]

Table 2 Temporal feature calculation components

<table>
<thead>
<tr>
<th>Function</th>
<th>mean, maximum, minimum, standard deviation, z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window Length</td>
<td>1 s, 2 s, 4 s</td>
</tr>
<tr>
<td>Data</td>
<td>( F^\text{left-arm}_i, F^\text{right-arm}_i, F^\text{left-leg}_i, F^\text{right-leg}_i )</td>
</tr>
<tr>
<td></td>
<td>( F^2, F^5, F^6, F^8, F^9, F^{10} )</td>
</tr>
</tbody>
</table>
2.5 Modeling Prechtl Cramped-Synchronized General Movements. The target classifications that we modeled were taken from the start and stop times that the trained physical therapist marked when observing CSGMs. The feature vector for each data point consisted of a total of 166 features, derived from both instantaneous data and temporal features as described above. The physical therapist’s annotations provided a binary target class for each feature vector that was the value that we were trying to model and predict.

To predict the class we utilized three statistical machine learning techniques: support vector machines (SVMs), decision trees, and dynamic Bayesian networks observing the output from random forests [37]. Each technique utilized cross-validation on a baby-by-baby basis in which a model was trained on one baby and tested on the other nine babies’ data. During testing, the first two techniques are presented with feature vectors in isolation. The DBN is presented with the complete temporal series of feature vectors for each training set. In all cases, the test data is withheld while training. Each algorithm chooses a classification for each sample based on its respective algorithm.

3 Results

In order to validate our experimental setup and the accelerometers’ ability to perform movement assessments, we first conducted an analysis in which a nurse scored the videos for periods of infant activity. Using the four-point Giganti scale [20], we had a nurse annotate the video data at 110 time points throughout the hour of data. Fifty annotations were made at the top of every minute for 50 min, and 60 annotations were made every ten seconds for the remaining 10 min. Visual and statistical analysis demonstrated high correlation between the nurse’s labels and the raw data. A portion of the data is shown in Fig. 2. Ten minutes of data from the four limbs are shown in the top four graphs black with the nurse’s scoring in red in the bottom graph.

The average overall corrected acceleration magnitude for all infants was 0.095 m/s\(^2\), with an average corrected arm acceleration of 0.11 m/s\(^2\) and average corrected leg acceleration of 0.08 m/s\(^2\). The maximum arm acceleration noted in any infant was 3.80 m/s\(^2\), and the maximum leg acceleration was 3.87 m/s\(^2\). Figure 3 shows the average corrected acceleration for each baby over the entire 1 h observation period (average across all times of the average of four limbs at each time sample). A wide range of average accelerations reflected wide variability in baby motion. Figure 4 shows the Pearson correlation calculated between the arms and legs.
between the legs of each baby. The right arm and left arm had an average correlation coefficient of 0.47, with 1 being perfectly correlated and 0 being no correlation. The right leg and left leg had an average correlation coefficient of 0.48. Figures 5 and 6 show a scatter plot of the arm and leg acceleration observed and the average corrected acceleration observed. A strong correlation is seen between arm and leg acceleration.

### 3.1 Prechtl Movement Results

The physical therapist identified 102 total CSGMs in six of the ten infants in the study. Each baby generated approximately 70,000 data points for a total of approximately 700,000 samples in our data. Using tenfold cross-validation, we applied the three algorithms to the data. Each technique produced different results that traded off sensitivity and specificity. Accuracies ranged from 70-90% and achieved an average sensitivity of 99.2% and an average specificity of 99.6%. Detailed measurements are shown in Table 3.

We conducted an analysis of the most informative features that were identified by the decision tree modeling. For each decision tree that was created in each of the ten cross validation folds, we ranked the features according to their independent information gain to identify which single feature was the most correlated with CSGMs. For each feature we averaged the rank across the ten trees. Table 4 shows the results.

The most informative feature was the minimum across a 2 s window of the maximum acceleration magnitude of all limbs for a given sample. This demonstrates that recognizing CSGMs requires observing a sustained motion. CSGMs are not particularly high-energy motions, however, and this suggests that normal motions do not sustain a continuously observed acceleration on all limbs for an entire 2 s window. The next nine most informative features were based on recognizing peak accelerations. This suggests that the sensor signature of CSGMs has sudden stops and starts although the movement itself is not characterized by a lot of high frequency motion. Finally, two of the top four features were correlated with the right side of the body. This is a curious result, but retrospective analysis of the cohorts medical records indicates that the majority of the patients had left side hemorrhaging that was apparently manifesting in right side motion.

### 4 Discussion

The diagnosis of neuromotor abnormalities in children, as well as the assessment of patterns of physical activity early in life, has traditionally rested on observation by physicians, skilled health care professionals, or parents [38,39]. Advances have been made in a variety of such approaches including tools such as the Test of Infant Motor Performance [40], the Alberta Infant Motor Scale [41], and the Infant Motor Profile [42]. However, these approaches are tedious, rely on skilled observers, require substantial time for observation and analysis, and are subject to substantial errors due to inter-rater reliability. Moreover, direct observations of human energy expenditure have been found to be limited to 3 s intervals of activity and to no more than 20 min of direct observation (before observer fatigue becomes problematic) [43]. Given the rapid and sporadic nature of physical activity in premature babies, direct observation of physical activity profiles has serious limitations.

In this study, we were able to successfully implement infant accelerometer monitoring in the NICU. We were able to...
Table 4  Average ranking of features according to information gain

<table>
<thead>
<tr>
<th>Feature Description</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>min (all limbs)</td>
<td>2.17</td>
</tr>
<tr>
<td>max (right arm)</td>
<td>3.67</td>
</tr>
<tr>
<td>max (left arm)</td>
<td>2.17</td>
</tr>
<tr>
<td>max (right leg)</td>
<td>4.25</td>
</tr>
<tr>
<td>max (left leg)</td>
<td>4.5</td>
</tr>
<tr>
<td>max (max both legs)</td>
<td>4.67</td>
</tr>
<tr>
<td>max (max both limbs)</td>
<td>5.5</td>
</tr>
<tr>
<td>max (max (both limbs))</td>
<td>6.5</td>
</tr>
<tr>
<td>max (left arm)</td>
<td>7.3</td>
</tr>
<tr>
<td>min (max both arms)</td>
<td>10.5</td>
</tr>
</tbody>
</table>

References


Acknowledgment

The authors thank Maria Coussens, RNC, Susan Gallitto, RNC, BSN who screened participants and assisted in accelerometer monitoring. This research was funded by National Institute for Nursing Research Grant No. RO1 NR09070 and by the UC Irvine Institute for Clinical and Translational Science (CTSA UL1RR031985).


