

Med Sci Sports Exerc. Author manuscript; available in PMC 2017 May 01.

Published in final edited form as:

Med Sci Sports Exerc. 2016 May; 48(5): 958–966. doi:10.1249/MSS.0000000000000842.

Decision Trees for Detection of Activity Intensity in Youth with Cerebral Palsy

Stewart G. Trost¹, Maria Fragala-Pinkham², Nancy Lennon³, and Margaret E. O'Neil⁴

¹Institute of Health and Biomedical Innovation @ the Queensland Centre for Children's Health Research, School of Exercise and Nutrition Sciences, Queensland University of Technology, Brisbane, QLD, Australia

²Research Center, Franciscan Hospital for Children, Brighton, MA

³Gait Analysis Laboratory, AI duPont Hospital for Children, Wilmington, DE

⁴Department of Physical Therapy and Rehabilitation Sciences, Drexel University, Philadelphia, PA

Abstract

Purpose—To develop and test decision tree (DT) models to classify physical activity (PA) intensity from accelerometer output and Gross Motor Function Classification System (GMFCS) classification level in ambulatory youth with cerebral palsy (CP); and 2) compare the classification accuracy of the new DT models to that achieved by previously published cut-points for youth with

Methods—Youth with CP (GMFCS Levels I - III) (N=51) completed seven activity trials with increasing PA intensity while wearing a portable metabolic system and ActiGraph GT3X accelerometers. DT models were used to identify vertical axis (VA) and vector magnitude (VM) count thresholds corresponding to sedentary (SED) (< 1.5 METs), light PA (LPA) (1.5 and <3 METs) and moderate-to-vigorous PA (MVPA) (3 METs). Models were trained and crossvalidated using the "rpart" and "caret" packages within R.

Results—For the VA (VA_DT) and VM decision trees (VM_DT), a single threshold differentiated LPA from SED, while the threshold for differentiating MVPA from LPA decreased as the level of impairment increased. The average cross-validation accuracy for the VC DT was 81.1%, 76.7%, and 82.9% for GMFCS levels I, II, and III, respectively. The corresponding crossvalidation accuracy for the VM_DT was 80.5%, 75.6%, and 84.2%, respectively. Within each GMFCS level, the DT models achieved better PA intensity recognition than previously published cut-points. The accuracy differential was greatest among GMFCS level III participants, in whom the previously published cut-points misclassified 40% of the MVPA activity trials.

Conclusion—GMFCS-specific cut-points provide more accurate assessments of MVPA levels in youth with CP across the full spectrum of ambulatory ability.

Keywords

Accelerometer; Motor Impairment; Children; Sedentary Behavior; Disability

Introduction

Cerebral palsy (CP) is a disorder of movement and posture secondary to an insult to the developing brain (26). It is the most common physical disability of childhood with a prevalence of 2.5 to 3.6 cases per 1000 live births (39). Inadequate physical activity (PA) and poor fitness are major problems affecting the health and well-being of children with CP (11). Furthermore, lack of PA may contribute to the development of disabling secondary conditions such as obesity, chronic pain, fatigue, and osteoporosis (11,17).

In recent years, strategies and goals for physical therapy and rehabilitation services for youth with CP have shifted from a focus on developmental motor skills and quality of movement to activity-based interventions for improved fitness, functional mobility and habitual PA (7,27,35). This shift is partially due to evidence suggesting that ambulant youth with CP experience decreased fitness and PA participation compared to age matched peers, which places them at increased risk for poor health outcomes later in life (35,36). Although therapeutic interventions typically include strategies to increase PA performance, researchers and clinicians have primarily relied on tests of functional capacity or self-report measures of PA to evaluate intervention intensity and effectiveness (36). While self-report measures are low-cost and convenient for participants to complete, they are subject to social desirability and considerable recall bias; and thus may not be sufficiently valid or reliable for clinical outcome studies involving youth with CP (30).

Accelerometry-based motion sensors have become the method of choice for assessing PA in children and adolescents (29). However, despite their widespread use among youth with typical development, calibrating these devices to units of energy expenditure or PA intensity in youth with CP presents significant methodological challenges. The atypical gait patterns and lower mechanical efficiency of children and adolescents with CP mandates that algorithms to delineate PA intensity be specifically developed for the CP population (5). Moreover, children and adolescents with CP exhibit significant heterogeneity with respect to functional severity which impacts both the accelerometer output and the energy cost of movement (17,33). Ambulant youth with CP are classified on the Gross Motor Function Classification System (GMFCS) at levels I – III (23). Youth at GMFCS level I have minimal motor impairment and activity limitation, while those at GFMCS II or III have moderate motor impairment and activity limitations (23). Children at GMFCS III use assistive devices (e.g., crutches, walkers) for ambulation and are at highest risk for adopting 'early inactivity' and becoming wheelchair users in early adulthood (8,23).

To date, a small number of studies have evaluated the validity of accelerometer-based motion sensors in children and adolescents with CP. The results of these studies show accelerometer output to be strongly correlated with measured energy expenditure and sensitive to changes in PA intensity (4,21,22). In addition, several investigators have completed calibration studies for the purpose of deriving intensity-related thresholds or cut-

points for classifying accelerometer counts as sedentary (SED), light-intensity PA (LPA), or moderate-to-vigorous intensity PA (MVPA). Clanchy et al. (6) derived intensity based cutpoints for the ActiGraph accelerometer using walking trial data from 29 ambulant children and adolescents with CP. The authors identified a VA count threshold of 503 counts per 15 sec for differentiating MVPA from LPA, which approximated the 573 counts per 15 sec threshold identified by Evenson and colleagues (10) in typically developing children. More recently, Oftedal et al. (20) derived an ActiGraph cut-point for differentiating non-sedentary from sedentary activity in toddlers with CP (mean age = 2.0 ± 0.5 y). The optimal threshold for the VM was 120 counts per 15 sec.

Although the aforementioned calibration studies have enabled researchers and clinicians to broadly quantify PA performance in children and adolescents with CP, the results are subject to a number of methodological limitations. First, although the Clanchy study (6) included participants across the full spectrum of ambulatory ability, the number of participants classified at Level III on the GMFCS scale was very small (N=3), raising questions about the validity of the cut-points among ambulant youth with more severe motor impairments. Second, the activities included in previous calibration studies have been limited to self-paced walking trials or therapist directed activities included in the clinical assessment of gross motor function. To date, no calibration study has derived cut-points using a range of real world or "lifestyle" activities such as sitting doing homework or completing household chores. Third, previously derived ActiGraph cut-points for MVPA have been based solely on VA counts. To date, no study has derived CP-specific cut-points for MVPA based on the VM of counts recorded in three planes. Finally, previous studies have used Receiver Operating Characteristic (ROC) Curves to identify intensity-based count thresholds. Because ROC Curves are based on binary outcomes (i.e., SED vs. other), a series of ROC Curve analyses are completed to identify the thresholds corresponding to SED, LPA, and MVPA. To our knowledge, no studies have implemented pattern recognition or machine learning approaches such as DT models to identify intensity-based count thresholds for children and adolescents with CP. DT models are well-suited to the task of deriving accelerometer cutpoints because they use entropy-based information gain indices to identify "splits" or count thresholds that maximize the classification of PA intensity. Moreover, decision trees can accommodate both continuous and categorical inputs and can perform multi-class classification tasks with a single model.

To address the aforementioned gaps in the research literature, the aims of the current study were 2-fold: 1) use DT models to identify intensity-based count thresholds for children and adolescents with CP across the full spectrum of ambulatory; and 2) compare the classification accuracy of the DT models to that achieved by previously published cut-points for children and adolescents with CP.

Methods

Participants

A total of 57 ambulatory youth with CP classified at GMFCS Levels I to III participated in the study. Participants were recruited from two pediatric hospitals through outpatient clinics and recruitment flyers distributed to clinicians in nearby hospitals. Children were ineligible

if they had: 1) undergone orthopedic surgery within the past 6 months; 2) lower extremity botulinum toxin injections within the last 3 months; or 3) a recent musculoskeletal injury or medical condition limiting their ability to complete the PA protocol. After deletions for missing energy expenditure and/or accelerometer data, the analytic sample comprised 51 participants. Of this number, 27 were classified at GMFCS Level I, 12 at GMFCS Level II, and 12 at GMFCS Level III. Descriptive characteristics for each GMFCS group are displayed in Table 1. The study was approved by the Institutional Review Boards at Franciscan Hospital for Children, Nemours/Alfred I. duPont Hospital for Children, and Drexel University. Prior to participation, parents and children provided written consent and assent.

Activity Protocol

Participants completed seven standardized activity trials while wearing a light-weight portable metabolic system (925 g) and tri-axial accelerometers on the left and right hip. The activity trials were completed in a single 2 h testing session and comprised the following sequence of activities: 1) Supine rest – lying down resting but not sleeping; 2) Seated handwriting – while sitting in a chair at a desk, using a pencil or pen to transcribe a standardized written script to a pad of paper; 3) Wiping down a countertop - walking from side to side in front of a 2 m long countertop spraying and wiping to clean the entire surface; 4) Folding laundry – loading a laundry bag with five towels and carry it 3 m; dumping out the towels, folding them, loading them back in the bag, carrying it back to the original starting spot, and repeating; 5) Comfortable walk - walking at comfortable self-selected speed after receiving the instructions "Walk at a comfortable pace like when you are at the mall or walking in your neighbourhood or at school but not in a hurry"; 6) Brisk walk walking at a brisk self-selected speed after receiving the instructions "This time we want you to walk a little faster so please walk at a faster pace like when you are hurrying to get to class after the bell has rung"; 7) Fast walk - walking at a fast speed after receiving the instructions "This time we want you to walk as fast as you possibly can without falling or running." By design, the trials were ordered by increasing intensity. This prevented physiological carry over effects from the previous trial and ensured that participants functioning at GMFCS Levels II or III would not be excessively fatigued before completing the walking trials. Trials 1 to 4 were 5 minutes in duration, while trials 5 to 7 were 6 minutes in duration. Between each trial, participants rested in a seated position until their HR returned to resting level. All walking trials were completed on a 25 m course marked with two cones. Prior to each walking trial, participants completed a brief practice test to help them select an appropriate waking speed. Additionally, during each walking trial, a physical therapist walked alongside each participant to help them maintain a consistent pace throughout the 6 minutes. The walking protocol was validated in a previous study (21,22).

Instrumentation

Portable Indirect Calorimetry—Oxygen uptake (VO₂) during each activity trial was measured using the Cosmed K4b² portable indirect calorimetry system (Rome, Italy). A flexible facemask (Hans Rudolf, Kansas City, MO) held in place by a nylon head harness covered the participants nose and mouth. The mask was attached to a bidirectional digital turbine flowmeter to measure the volume of inspired and expired air. A sample line running

from the turbine to the analyzer unit delivered expired air for the determination of O_2 and CO_2 content. The Cosmed K4b2 has been shown to provide valid measures of oxygen uptake over a range of exercise intensities (24) and has been validated for youth with CP (21,22).

Accelerometry—The ActiGraph GT3X (ActiGraph Corporation, Pensacola, FL) measures acceleration in the vertical [x], medio-lateral [y], and antero-posterior [z] planes ranging in magnitude from 0.05-2.5 g at a sampling rate of 30 Hz. Once digitized, the output signal passes through a filter that band limits the device to a frequency range of 0.25-2.5 Hz. The filtered signal is then rectified and integrated over a user-specified interval known as an epoch. At the end of each epoch, the summed value or "activity count" is stored in memory and the integrator is reset. To facilitate synchronization with the indirect calorimetry data, a 1 second epoch was utilized. Before each testing session, the accelerometers were initialized according to the manufacturer's specifications and attached to a flexible elastic belt that was fastened snugly around the waist of the participant. The accelerometers were positioned on the left and right mid-axilla line at the level of the iliac crest. Given the previously reported high level of agreement between monitors placed on the left and right hips observed in this sample (22), only accelerometer output from the right hip were used in the current analyses.

Data Reduction

Prior to each testing session, the accelerometer and portable indirect calorimetry units were synchronized to an external timepiece. After download of the data from the respective instruments, customized software was used to reintegrate the accelerometer data to 15 s epochs (to match previously published cut-points) and calculate mean VO₂, mean vertical axis (VA) counts and mean vector magnitude (VM) [(x0221a) $(x^2 + y^2 + z^2)$] from the middle portion of each activity trial. For the 5-min activity trials, data from minutes 2 to 4 were averaged. For each of the 6-minute walking trials, data from minutes 2 to 5 were averaged. Data from the middle portion of each trial were used to ensure that participants were at a steady pace and not building up speed at the start of the trial or slowing down in anticipation of the end of the trial. For each activity trial, energy expenditure in METs was calculated by dividing activity VO₂ by resting VO₂, where resting VO₂ was predicted using Schofield's equations (28). Based on these MET values, activity trials were classified as SED, LPA, or MVPA. SED was defined as <1.5 METs, LPA was defined as 1.5 and <3 METs; and MVPA was defined as 3 METs (22,35). To evaluate the classification accuracy of the Clanchy (CL) and Evenson (EV) cut-points, mean VA counts for each trial was classified as SED, LPA, and MVPA using the following thresholds – EV: SED < 25 cnts per 15 s, MVPA 574 cnts per 15 s; CL: SED < 25 cnts per 15 s, MVPA 503 cnts per 15 s.

Classification Trees

Binary DT models were used to identify new VA and VM count thresholds corresponding to SED, LPA, and MVPA. The DT starts with a single root node that splits into multiple branches, leading to further nodes, each of which may be further split or terminate as a leaf node. Each non-leaf node is associated with a specific rule or question which determines which branch to follow. The DT is built by identifying the variable (in this case - accelerometer counts) that splits the training data into subgroups with the greatest class label

purity (i.e., only one level of physical activity intensity). After each split, the data instances are further divided into nodes defined by the splitting variable, and splitting continues in each node until the subgroups either reach a minimum size or until some stop condition is reached. In the current study, splits were based on the information gain index which identifies the split point providing the greatest reduction in entropy or label impurity.

In the current study, DT models were trained, tuned, and cross-validated using the "rpart" and "caret" packages within R (15). The "rpart" package implements the Classification and Regression Trees (CART) algorithm of Breiman and colleagues (3). Separate tree models were developed for each GMFCS level with VA or VM counts serving as the only input. The minimum number of instances allowed in a node prior to a split was set to 20, while the minimum number of instances allowed in any one leaf node was set to 7.

The train function within the "caret" package was used to tune candidate values of the complexity parameter and evaluate model performance using leave-one-out cross-validation (LOOCV). The complexity parameter controls the process of pruning a DT. It is used to control the size of the DT and to select an optimal tree size. For the current study, accuracy was optimized at a complexity parameter value of 0.1. In LOOCV, the prediction model is trained on data from all of the participants except one, which is "held out" and used as the test dataset. The process is repeated until all participants have served as the test data set and the accuracy results are averaged.

Model Evaluation

For the previously published and newly derived cut-points, classification accuracy was evaluated by calculating weighted Kappa statistics, sensitivity (Se), specificity (Sp), and area under the Receiver Operating Characteristic curve (ROC-AUC). For interpretation of the Kappa coefficients (κ), we followed the ratings suggested by Landis and Koch (16): poor (0–0.2), fair (0.2–0.4), moderate (0.4–0.6), substantial (0.6–0.8), and almost perfect (0.8–1.0). For ROC-AUC, an area of 1.0 represents perfect classification, whereas an area of 0.5 represents an absence of classification accuracy. ROC-AUC values of 0.90 are considered excellent, 0.80–0.89 good, 0.70–0.79 fair, and < 0.70 poor (19).

Results

Table 2 displays descriptive statistics for energy expenditure (METs), VA counts, and VM counts for the seven activity trials by GMFCS level. Within GMFCS level, MET values and accelerometer output increased in a dose-response manner as the required work rate for each trial increased. Across GMFCS levels, MET values and accelerometer output varied considerably. For the sedentary activity trials, METs and accelerometer output were comparable across GMFCS level. For light-intensity household chores, MET levels were comparable across GMFCS level; however, accelerometer output tended to decrease as GMFCS level increased from I to III. During the comfortable and brisk paced walking trials, MET values increased with GMFCS level, while accelerometer output decreased. During fast-paced walking, GMFCS III participants exhibited substantially lower VA and VM outputs relative to their more ambulatory counterparts, despite exhibiting comparable MET values. For GMFCS I participants, average walking speed during the comfortable, brisk, and

fast-paced walking trials was 52.6 ± 3.1 , 69.5 ± 4.2 , and 87.5 ± 5.2 m/min, respectively. For GMFCS II participants, average walking speed was 42.1 ± 2.5 , 58.0 ± 3.5 , and 69.3 ± 4.2 m/min, respectively. For GMFCS III participants, average walking speed was 26.8 ± 1.6 , 42.0 ± 2.5 , and 56.9 ± 3.4 m/min, respectively.

The aggregated DT model for VA counts (VA_DT) and the VM (VM_DT) are displayed in Figures 1 and 2, respectively. In both models, a single threshold differentiated SED from LPA (VA < 8 cnts per 15 s, VM < 72 cnts per 15 s), while the threshold for differentiating LPA from MVPA decreased as the level of impairment increased (VA: GMFCS I < 535 cnts per 15 s, GMFCS II < 333 cnts per 15 s, GMFCS III < 200 cnts per 15 s; VM: GMFCS I < 724 cnts per 15 s, GMFCS II < 685 cnts per 15 s, GMFCS III < 669 cnts per 15 s). In LOOCV, the average cross-validation accuracy for the VA_DT was 81.1% (95% CI = 74.7 – 86.5), 76.7% (95% CI = 66.4 – 85.2), and 82.9% (95% CI = 73.0 – 90.3) for youth in GMFCS levels I, II, and III, respectively. The corresponding cross-validation accuracy for the VM_DT was 80.5% (95% CI = 74.1 – 86.0), 75.6% (95% CI = 65.1 – 84.2), and 84.2% (95% CI = 74.4 – 91.3), respectively. Aggregated across all three GMFCS levels, the average cross-validation accuracy for the VA_DT and VM_DT was 80.5% (95% CI = 76.3 – 84.6) and 80.2% (95% CI = 76.0 – 84.3), respectively.

Table 3 reports Se, Sp, and ROC_AUC statistics for the previously established cut-points and the newly derived DT models. For SED, the EV and CL cut-points exhibited excellent classification accuracy, with Se, Sp, and ROC-AUC exceeding 90%. The VA_DT and VM_DT models also exhibited excellent classification accuracy, but demonstrated better Sp than the EV and CL cut-point. For LPA, accuracy for the EV and CL cut-points was poor. The poor classification accuracy was a function of both low Se and Sp. In comparison, classification accuracy for the VA_DT and VM_DT was good, with both models providing higher Se and Sp than the EV and CL cut-points. For MVPA, the CP and EV cut-points exhibited fair classification accuracy. These cut-points were characterised by high Sp, but low Se. In comparison, classification accuracy for the VA_DT and VM_DT was good, with both models providing higher Se and ROC-AUC than the EV and CL cut-points.

Applying the rubric of Landis and Koch (16), the EV (κ = 0.67, 95% C.I. = 0.61 – 0.72) and CL (κ = 0.65, 95% C.I. = 0.60 – 0.72) cut-points exhibited moderate to substantial agreement, while the VA_DT (κ = 0.81, 95% C.I. = 0.77 – 0.86) and VM_DT (κ = 0.80, 95% C.I. = 0.76 – 0.85) models exhibited substantial to almost perfect agreement across all three levels of PA intensity.

Figure 3 displays GMFCS-specific misclassification rates for the two DT models relative to the EV and CL cut-points. The misclassification rate is equal to the sum of the false negatives and false positives divided by total number of cases. For all four models, the misclassification rate for SED was relatively low (< 12%). Nonetheless, SED misclassification rates for the VA_DT and VM_DT (1.2% - 7.3%) were lower than that observed for the EV and CL cut-points (8.1% - 11.6%). For LPA, misclassification rates for the VA_DT and VM_DT (11.0% - 18.6%) were substantially lower than that observed for the EV and CL cut-points (23.2% - 43.9%), particularly among GMFCS III participants, in whom the EV and CL cut-points misclassified over 40% of the LPA trials. Among GMFCS

III participants, MVPA misclassification rates for the VA_DT (9.8%) and VM_DT (11.0%) were approximately one-third lower than that observed for the EV and CL cut-points (31.7% - 34.1%). Among GMFCS I and II participants, the VA_DT (14.0% – 16.2%) exhibited similar or marginally lower MVPA misclassification rates than the EV and CL cut-points (16.2% - 18.6%).

Discussion

The current study is the first to use binary DT models to identify and validate ActiGraph cutpoints for youth with CP across the full spectrum of ambulatory ability. The resultant models, which identified GMFCS-specific thresholds for MVPA, achieved greater than 80% accuracy in the LOOCV and exhibited better Se, Sp, and overall classification accuracy than previously published cut-points. Within each GMFCS level, the newly derived DT models achieved better PA intensity recognition than the CL and EV cut-points. The accuracy differential was greatest among GMFCS level III participants, among whom the previously published cut-points misclassified 40% of the LPA trials and 30% of the MVPA activity trials.

The progressively lower MVPA cut-points identified for youth in GMFCS Levels I through III is consistent with the results of previous research reporting a strong positive correlation between GMFCS level and the energy cost of locomotion (14,25). Although participants with more severe motor impairments completed the self-paced walking trials at a slower pace and generally accumulated fewer accelerometer counts per unit of time, the associated metabolic demand of performing these trials was sufficient to surpass the 3-MET MVPA threshold. In effect, there was "left shift" in the theoretical regression line summarizing the relationship between MET level and accelerometer counts. Thus, the magnitude of accelerometer output (VA or VM counts) associated with an absolute energy expenditure of 3 METs was, on average, lower among GMFCS Level II and III participants than those classified as GMFCS Level I. Of note, the new MVPA cut-point identified for GMFCS Level I participants was lower than the optimal cut-point reported for typically developing children (10), but higher than the MVPA threshold identified by Clanchy and colleagues (6) in a sample that featured a combination of youth in GMFCS Levels I and II.

Previous descriptive studies involving youth with CP have reported a large differential in PA performance across GMFCS level. Gorter and colleagues (12) used the CL cut-point to characterize daily MVPA patterns in ambulatory adolescents with CP. As GMFCS level progressed from Level I to III, daily MVPA participation declined from 56 min per day to just 9.3 min per day. Bjornson et al. (2) used the StepWatch monitor to characterize the daily MVPA levels of 81 youth with CP. Applying a generic step cadence threshold of 15 strides per minute, children in GMFCS level III were found to accumulate approximately 50% less time in PA than children in GMFCS levels I and II. In similar fashion, Van Wely et al. (34) assessed the ambulatory activity of 62 children with CP using the StepWatch monitor. Applying the same threshold of 15 strides per minute, children in GFMCS level III exhibited approximately 50% less time in moderate-to-high intensity PA than children in GMFCS level I and II. The results of the current study demonstrate that the relationship between absolute energy expenditure and accelerometer output varies considerably by GMFCS level.

Hence, the application of "one size fits all" cut-points developed for youth with typical development or those with low to mild functional impairment may dramatically underestimate the PA performance of youth with more severe motor impairments. This raises the possibility that the reported differences in MVPA across GMFCS levels may not be as large as previously thought. Future studies should test this intriguing hypothesis by applying the newly derived GMFCS specific cut-points in independent samples of free-living children and adolescents with CP.

A unique aspect of this study was the development of CP-specific MVPA cut-points based on the VM. Although we hypothesized that VM prediction models would provide more accurate assessments of PA intensity, the VA and VM DT models provided congruent classification accuracy. This finding is in conflict with the results of a recent study reporting a VM cut-point to be more accurate than a VA threshold in identifying bouts of nonsedentary activity in toddlers with CP (20). However, it is important to note that this study derived cut-points for differentiating non-sedentary from sedentary behaviour, and did not identify thresholds for LPA and MVPA. By recording acceleration in multiple planes, it is often assumed that tri-axial accelerometers are better able to quantify the variable movement patterns of free-living youth. However, both laboratory and field-based validation studies have failed to support this assumption (13,30). This finding may be explained, in large part, by the placement of the accelerometer on the body's centre of mass and the resultant strong correlation between VA and VM counts during upright ambulatory activity. Indeed, across the seven activity trials evaluated in the current study, the correlation between VA counts and the VM was r= 0.93. Thus, relative to VA counts, the VM provided little unique information about the intensity of PA. It may be that the advantages of collecting high frequency tri-axial accelerometer signal can only be fully realized if the numerous statistical and frequencydomain features within the data are extracted and inputted into more sophisticated pattern recognition or machine learning models (e.g., neural networks, support vector machines) (31,32). Accordingly, the application of machine learning approaches to tri-axial accelerometer data collected in youth with CP represents an important area of future methodological research.

In light of the recent interest in exploring interventions to decrease sedentary behaviour and promote light-intensity PA in youth with CP (35,37), the assessment of sedentary behaviour in this population has become a topic of considerable importance. The current study evaluated the predictive validity of the widely used sedentary cut-point of 100 counts per min, which previous research has shown to provide 92% classification accuracy among ambulatory youth with CP (6). Within our sample, the 100 counts per min cut-point (25 counts per 15 sec) provided excellent classification accuracy, with an ROC-AUC of 93%. Nevertheless, the newly derived DT models, which identified sedentary cut-points of 8 counts per 15 sec for VA counts and 72 counts per 15 sec for the VM, provided superior classification accuracy (ROC-AUC > 97%) and fewer misclassification errors within each GMFCS level. Importantly, the cut-points for delineating LPA from SED did not differ by GMFCS level. This result was not unexpected considering that sedentary behaviour accelerometer thresholds are based on the relative absence of movement, and are therefore less influenced by differences in ambulatory ability. It should be noted, however, that the DT models were developed using a purely intensity-based definition of sedentary behaviour (<

1.5 METs), which did not consider the participant's posture or the energetics of standing (37). The inability of waist- or hip-mounted accelerometers to reliably distinguish between the energy cost of sitting and standing is well-documented (29), and this limitation applies to the sedentary thresholds identified in the current study. Consequently, the development of more advanced measures of sedentary behaviour which consider both posture and energy expenditure is a priority for future CP-related research.

The current study had several limitations that warrant consideration. First, in order to obtain steady-state measures of energy expenditure that varied in intensity from SED to MVPA, the DT models were trained on data from controlled activity trials which do not fully replicate the activity patterns of free-living youth with CP. Second, due to the difficulty and participant burden associated with measuring resting energy expenditure in youth with CP, MET values were calculated by dividing measured VO2 values by predicted resting energy expenditure. Such an approach has been implemented in other calibration studies involving children with CP (6), and previous research has shown Schofield equations to provide valid predictions of resting metabolism in youth with CP (1). Third, because the selection of the complexity parameter was based on performance estimation in the LOOCV data, the reported error rates for both trees may be biased. Finally, the DT models were developed for processed activity counts calculated by ActiGraph's proprietary algorithm. Hence, the resultant cut-points are only generalizable to ActiGraph accelerometers and are not applicable to the processed counts or raw acceleration metrics provided by other makes and models of accelerometer-based motion sensors.

Offsetting limitation were several notable strengths. First, the study sample had sufficient numbers of youth in GMFCS levels II and III to evaluate and identify accelerometer cutpoints across the full spectrum of ambulatory ability. Second, unlike previous calibration studies which only included walking trials, the study protocol consisted of a mix of "lifestyle" and ambulatory activities ranging in intensity from sedentary to vigorous, thus adhering to the best practice recommendations outlined by Welk and others (38). Third, in contrast to previous calibration studies, which used ROC curve analyses to derive intensity-related thresholds, the present study used DT models and entropy-based information gain indices to identify optimal "splits" or cut-points for the recognition of SED, LPA, and MVPA. Because ROC Curves are based on binary outcomes (i.e., SED vs. other), multiple ROC Curves are needed to identify the thresholds corresponding to SED, LPA, and MVPA. Fourth and finally, the accuracy of the DT models were cross-validated using LOOCV which more closely simulates model performance when deployed as an "off the shelf" algorithm for use in independent samples of youth with CP (9).

In summary, previously published ActiGraph cut-points for children and adolescents with CP are not valid among youth with more severe motor impairments. DT models identifying GMFCS-specific thresholds for MVPA substantially reduced the misclassification of PA intensity among youth in GMFCS Levels II and III and provided more accurate assessments of PA performance across the full spectrum of ambulatory ability. Future studies should employ direct observation or other criterion measures of PA intensity to evaluate the predictive validity of the newly developed DT models in true free-living samples of ambulant youth with CP.

Acknowledgments

We thank the families for their participation and the research teams at the clinical sites (Franciscan Hospital for Children and Nemours duPont Hospital for Children). Funding for this project was provided by NIH R24 HD065688-01: Boston Rehabilitation Outcomes Center (Boston ROC); Improving Outcome Measurement for Medical Rehabilitation Clinical Trial. Professor Stewart Trost is a member of the ActiGraph Scientific Advisory Board. The results of the present study do not constitute endorsement by ACSM.

References

- Bandini LG, Schoeller DA, Fukagawa NK, Wykes LJ, Dietz WH. Body composition and energy expenditure in adolescents with cerebral palsy or myelodysplasia. Pediatr Res. 1991; 29(1):70–7. [PubMed: 2000262]
- 2. Bjornson KF, Belza B, Kartin D, Logsdon R, McLaughlin JF. Ambulatory physical activity performance in youth with cerebral palsy and youth who are developing typically. Phys Ther. 2007; 87(3):248–57. [PubMed: 17244693]
- Breiman, L.; Friedman, JH.; Olshen, RA.; Stone, CI. Classification and regression trees. Boca Raton, FL: CRC Press; 1984. p. 18-55.
- 4. Capio CM, Sit CH, Abernethy B. Physical activity measurement using MTI (Actigraph) among children with cerebral palsy. Arch Phys Med Rehabil. 2010; 91(8):1283–90. [PubMed: 20684912]
- Clanchy KM, Tweedy SM, Boyd RN. Measurement of habitual physical activity performance in adolescents with cerebral palsy: a systematic review. Dev Med Child Neurol. 2011; 53(6):499–505.
 [PubMed: 21418195]
- Clanchy KM, Tweedy SM, Boyd RN, Trost SG. Validity of accelerometry in ambulatory children and adolescents with cerebral palsy. Eur J Appl Physiol. 2011; 11(12):2951–9. [PubMed: 21442163]
- 7. Damiano DL. Activity, Activity, Activity: Rethinking our physical therapy approach to cerebral palsy. Phys Ther. 2006; 86(11):1534–40. [PubMed: 17094192]
- 8. Day S, Wu Y, Strauss D, Shavelle R, Reynolds R. Change in ambulatory ability of adolescents and young adults with cerebral palsy. Dev Med Child Neurol. 2007; 49(9):647–53. [PubMed: 17718819]
- 9. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiol Meas. 2014; 35(11):2191–2203. [PubMed: 25340969]
- Evenson KR, Cattellier D, Gill K, Ondrak K, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008; 26(14):1557–65. [PubMed: 18949660]
- Fowler E, Kolobe TH, Damiano D, et al. Promotion of physical fitness and prevention of secondary conditions for children with cerebral palsy: section on pediatrics research summit proceedings. Phys Ther. 2007; 87(11):1495–510. [PubMed: 17895351]
- 12. Gorter JW, Noorduyn SG, Obeid J, Timmons BW. Accelerometry: a feasible method to quantify physical activity in ambulatory and non-ambulatory adolescents with cerebral palsy. Int J Pediatr. 2012 Article ID 329284: 6 pages. 10.1155/2012/329284
- Howe CA, Staudenmayer JW, Freedson PS. Accelerometer prediction of energy expenditure: vector magnitude versus vertical axis. Med Sci Sports Exerc. 2009; 41(12):2199–206. [PubMed: 19915498]
- Johnston RE, Moore SE, Quinn LT, Smith BT. Energy cost of walking in children with cerebral palsy: relation to the Gross Motor Classification System. Dev Med Child Neurol. 2004; 46(1):34– 38. [PubMed: 14974645]
- 15. Kuhn M. Building predictive models in R using the caret package. J Stat Software. 2008; 28(5):1–26
- 16. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–74. [PubMed: 843571]

17. Maltais DB, Pierrynowski MR, Galea VA, Matsuzaka A, Bar-Or O. Habitual physical activity levels are associated with biomechanical walking economy in children with cerebral palsy. Am J Phys Med Rehab. 2005; 84(1):36–45.

- Maltais DB, Wiart L, Fowler E, Verschuren O, Damiano DL. Health-related physical fitness for children with cerebral palsy. J Child Neurol. 2014; 29(8):1091–1100. [PubMed: 24820339]
- 19. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978; 8:283-98. [PubMed: 112681]
- Oftedal S, Bell KL, Davies PS, Ware RS, Boyd RN. Validation of accelerometer cut points in toddlers with and without cerebral palsy. Med Sci Sports Exerc. 2014; 46(9):1808–15. [PubMed: 25134003]
- O'Neil ME, Fragala-Pinkham M, Forman J, Trost SG. Measuring reliability and validity of the ActiGraph GT3X accelerometer for children with cerebral palsy: a feasibility study. J Pediatr Rehabil Med. 2014; 7(3):233–40. [PubMed: 25260506]
- 22. O'Neil ME, Fragala-Pinkham M, Lennon N, George A, Forman J, Trost SG. Reliability and validity of objective measures of physical activity in youth with cerebral palsy who are ambulatory. Phys Ther. 2015 Jun 18. [Epub ahead of print].
- Palisano R, Rosenbaum P, Bartlett D, Livingston M. Content validity of the expanded and revised Gross Motor Function Classification System. Dev Med Child Neurol. 2008; 50(10):744–750.
 [PubMed: 18834387]
- Rosdahl H, Gullstrand L, Salier-Eriksson J, Johansson P, Schantz P. Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method. Eur J Appl Physiol. 2010; 109:159–71.
 [PubMed: 20043228]
- 25. Rose J, Gamble JG, Medeiros J, Burgos A, Haskell WL. Energy cost of walking in normal children and those with Cerebral Palsy: Comparison of heart rate and oxygen update. Int J Pediatr Orthop. 1989; 9(3):276–79.
- 26. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M. The definition and classification of cerebral palsy. Dev Med Child Neurol. 2007; 49(s109):8–14.
- 27. Rowland JL, Fragala-Pinkham M, Miles C, O'Neil ME. The scope of pediatric physical therapy practice in health promotion and fitness for youth with disabilities. Pediatr Phys Ther. 2015; 27(1): 2–15. [PubMed: 25521261]
- 28. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr. 1985; 39(1 suppl):5–41. [PubMed: 4044297]
- Trost SG. State of the art reviews: Measurement of physical activity in children and adolescents.
 Am J Lifestyle Med. 2007(1):299–314.
- 30. Trost SG, O'Neil, M. Clinical use of objective measures of physical activity. Br J Sports Med. 2014; 48(3):178–81. [PubMed: 24311601]
- 31. Trost SG, Wong W, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. Med Sci Sports Exerc. 2012; 44(9):1801–9. [PubMed: 22525766]
- 32. Trost SG, Zheng Y, Wong WK. Machine learning for activity recognition. Hip versus wrist data. Physiol Meas. 2014; 35(11):2183–9. [PubMed: 25340887]
- 33. van den Hecke A, Malghem C, Renders A, Detrembleur C, Palumbo S, Lejeune TM. Mechanical work, energetic cost, and gait efficiency in children with cerebral palsy. J Pediatr Orthop. 2007; 27(6):643–647. [PubMed: 17717464]
- 34. Van Wely L, Becher JG, Balesman ACJ, Dallmeijer AJ. Ambulatory activity of children with cerebral palsy: which characteristics are important? Dev Med Child Neuro. 2012; 54(5):436–42.
- 35. Verschuren O, Darrah J, Novak I, Ketelaar M, Wiatt L. Health enhancing physical activity in children with cerebral palsy: More of the same is not enough. Phys Ther. 2014; 94(2):297–305. [PubMed: 24092902]
- 36. Verschuren O, Ketelaar M, Gorter JW, Helders PJM, Uiterwaal CSPM, Takken T. Exercise training program in children and adolescents with cerebral palsy: A randomized controlled trial. Arch Pediatr Adolesc Med. 2007; 161(11):1075–81. [PubMed: 17984410]
- 37. Verschuren O, Peterson MD, Leferink S, Darrah J. Muscle activation and energy-requirements for varying postures in children and adolescents with cerebral palsy. J Pediatr. 2014; 165(5):1011–6. [PubMed: 25151195]

38. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med Sci Sports Exerc. 2005; 37(11 suppl):S501–11. [PubMed: 16294113]

39. Winter S, Autry A, Boyle C, Yeargin-Allsopp M. Trends in the prevalence of cerebral palsy in a population-based study. Pediatr. 2002; 110(6):1220–25.

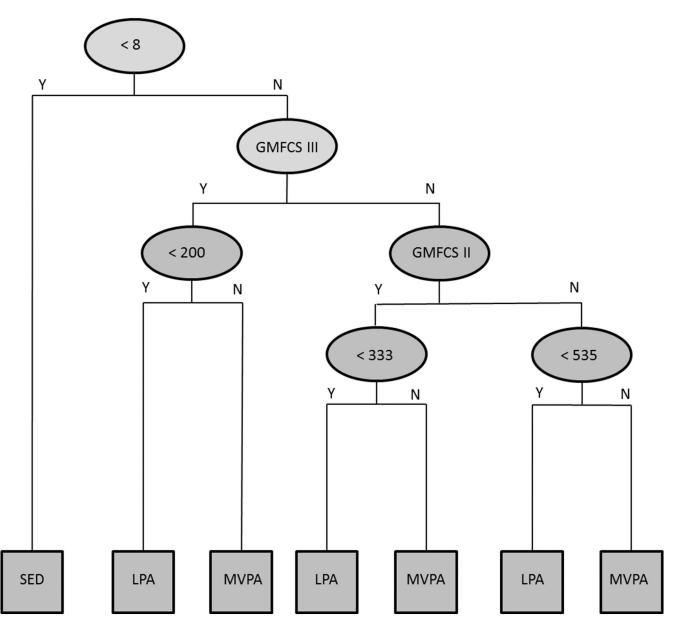


Figure 1. Decision tree model for identifying physical activity intensity from Gross Motor Function Classification System (GMFCS) level and ActiGraph vertical axis counts (counts per 15 s). Y = Yes, N = No, SED = Sedentary activity, LPA = Light intensity physical activity, MVPA = Moderate-to-vigorous physical activity.

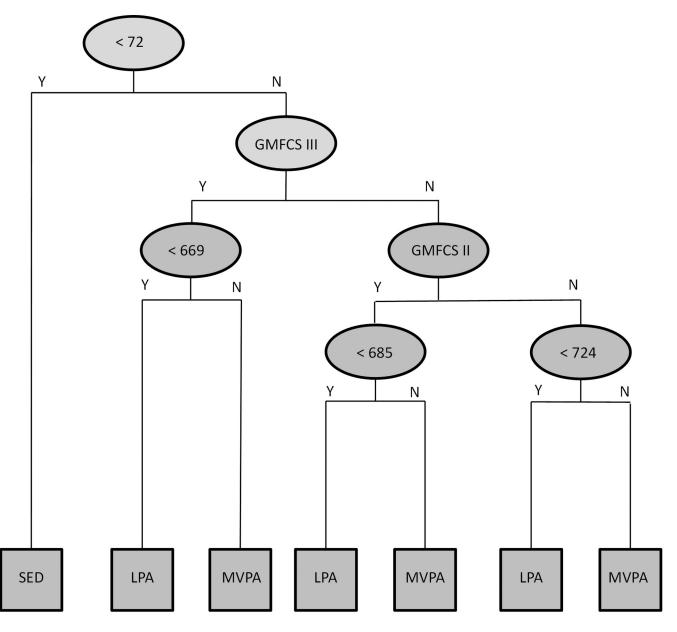


Figure 2.Decision tree model for identifying physical activity intensity from Gross Motor Function Classification System (GMFCS) level and ActiGraph vector magnitude (counts per 15 s). Y = Yes, N = No, SED = Sedentary activity, LPA = Light intensity physical activity, MVPA = Moderate-to-vigorous physical activity.

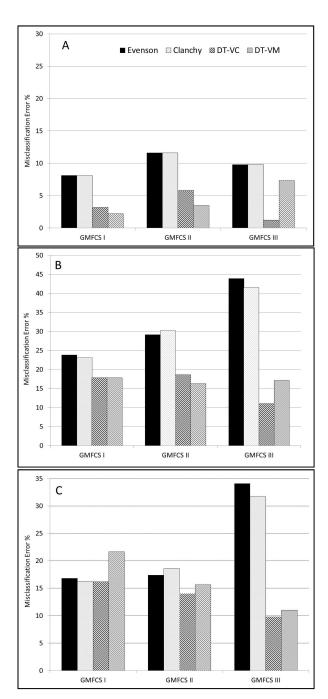


Figure 3. Misclassification rates for previously published Evenson and Clanchy cut-points and the newly develop vertical axis (VA_DT) and vector magnitude (VM_DT) decision tree models. Panel A: Sedentary activity, Panel B: light-intensity physical activity, Panel C: Moderate-to-vigorous physical activity (MVPA).

Table 1

Participants Characteristics

	GMFCS I (N=27)	GMFCS II (N=12)	GMFCS III (N=12)
Sex			
Male %	63	33	50
Female %	37	67	50
Diagnosis			
Hemiplegia %	74.1	33.3	8.3
Diplegia %	22.2	58.4	83.4
Quadriplegia %	3.7	8.3	8.3
Age (y)	12.4 ± 3.3	12.3 ± 3.4	12.7 ± 3.1
Height (cm)	147.1 ± 17.3	145.1 ± 15.3	147.0 ± 16.5
Weight (Kg)	42.8 ± 15.9	43.2 ± 11.9	46.8 ± 19.0
$Resting\ VO_{2}\ (\text{ml/kg/min})$	4.6 ± 1.0	4.5 ± 0.9	4.5 ± 0.9

Author Manuscript

Author Manuscript

Descriptive statistics for energy expenditure (METs) and accelerometer output by activity trial and GMFCS level Table 2

	MET	METs(Mean & 95% C.L.)	CL)	Vertical Axis Counts per 15 s (VA)(Median & 95% C.I.)	per 15 s (VA)(Med	ian & 95% C.L.)	Vector Magnitud	Vector Magnitude per 15 s (VM)(Median & 95% C.I.)	an & 95% C.I.)
	GMFCS I	GMFCS I GMFCS II	GMFCS III	GMFCS I	GMFCS II	GMFCS III	GMFCS I	GMFCS II	GMFCS III
Lying Down	1.1(1.0 – 1.2)	1.1(1.0-1.2) $1.2(1.1-1.4)$	1.0(0.9 - 1.2)	0(0-0)	0(0-1)	0(0-0)	0(0-0)	0(0-1)	0(0-0)
Handwriting	1.2(1.1 - 1.3)	1.2(1.1-1.3) $1.3(1.1-1.4)$	1.1(0.9 - 1.2)	0(0-0)	0(0-2)	(0-0)0	1(0-7)	8(1-22)	1(0-2)
Clean Counter	2.9(2.6 - 3.2)	2.9(2.6 - 3.2) $2.8(2.5 - 3.0)$	2.7(2.4 - 3.0)	96(64 - 136)	58(36 - 126)	61(43 - 110)	601(549 - 658)	546(385 - 645)	338(233 – 442)
Laundry Task	2.4(2.1 - 2.7)	2.4(2.1-2.7) $2.4(2.2-2.7)$	2.3(2.0 - 2.7)	35(26-71)	26(14-45)	15(9-35)	352(293 - 390)	280(144 - 371)	138(71 - 211)
Comfortable Walk	2.8(2.6 – 3.1) 3.3(2.9 – 3.7)	3.3(2.9 - 3.7)	3.5(3.0 - 4.0)	450(382 - 547)	521(415 – 755)	276(199 – 377)	793(672 – 922)	942(894 - 1026)	720(670 - 1052)
Brisk Walk	3.7(3.3 – 4.1)	3.7(3.3 - 4.1) $4.0(3.5 - 4.5)$	4.3(3.7 - 4.8)	890(718 – 995)	946(804 - 1018)	450(294 – 643)	1191(1026 - 1287)	1311(1169 – 1520)	958(768 - 1377)
Fast Walk	4.7(4.1 – 5.3)	4.7(4.1 - 5.3) $4.9(4.2 - 5.5)$	4.7(4.1 - 5.2)	1200(1091 - 1287)	1017(898 - 1108)	587(341 – 903)	1460(1331 - 1662)	1429(1285 - 1554)	1152(757 - 1533)

Author Manuscript

Author Manuscript

Classification accuracy for previously published cut-points and the newly developed decision tree models Table 3

	Evenson	et al. (95% CI)	Clanchy	et al. (95% CI)	Decision T	Evenson et al. (95% CI) Clanchy et al. (95% CI) Decision Tree (VA) (95% CI) Decision Tree (VM) (95% CI)	Decision T	ree (VM) (95% CI)
SED								
%e%	6.86	(96.9 - 100)	6.86	(96.9 - 100)	6.76	(95.1 - 100)	6.96	(93.4 - 100)
%dS	9.78	(83.5 - 91.6)	87.6	(83.5 - 91.6)	96.1	(93.7 - 98.5)	96.1	(93.7 - 98.5)
ROC-AUC	0.93	(0.91 - 0.96)	0.93	(0.91 - 0.96)	0.97	(0.95 - 0.99)	96.0	(0.94 - 0.99)
LPA	25 cm	25 cnts per 15 sec	25 cm	25 cnts per 15 sec	8 cm	8 cnts per 15 sec	72 cr	72 cnts per 15 sec
%e%	61.5	(52.7 - 70.4)	58.1	(49.2 - 67.1)	77.8	(70.2 - 85.3)	72.7	(64.6 - 80.7)
%dS	74.6	(69.0 - 80.1)	77.1	(71.8 - 82.5)	86.4	(82.1 - 90.8)	87.8	(83.5 - 91.9)
ROC-AUC	89.0	(0.63 - 0.73)	89.0	(0.62 - 0.73)	0.82	(0.77 - 0.86)	0.80	(0.76 - 0.85)
MVPA	574 ca	574 cnts per 15 sec	503 cr	503 cnts per 15 sec	GMFC	GMFCS Level Specific	GMFC	GMFCS Level Specific
%e%	57.1	(48.9 - 65.5)	61.4	(53.4 - 69.5)	78.6	(71.8 - 85.4)	81.4	(75.0 - 87.9)
%dS	93.4	(90.1 - 96.8)	91.6	(87.8 - 95.3)	92.5	(89.0 - 96.0)	7.68	(85.6 - 93.8)
ROC-AUC	0.75	(0.71 - 0.80)	0.76	(0.72 - 0.81)	0.86	(0.82 - 0.89)	0.86	(0.82 - 0.89)

SED = Sedentary, LPA = Light Physical Activity, MVPA = Moderate-to-Vigorous Physical Activity, Se% = Sensitivity, Sp% = Specificity, ROC-AUC = Area under the Receive Operator Characteristic (ROC) Curve