

Validation of Accelerometer Prediction Equations in Children with Chronic Disease

Abstract:

Purpose: To assess the criterion validity of existing accelerometer-based energy expenditure (EE) prediction equations among children with chronic conditions, and to develop new prediction equations.

Methods: Children with congenital heart disease (CHD), cystic fibrosis (CF), dermatomyositis (JDM), juvenile arthritis (JA), inherited muscle disease (IMD), and haemophilia (HE), completed 7 tasks while EE was measured using indirect calorimetry with counts determined by accelerometer. Agreement between EE predicted and measured EE was assessed. Disease-specific equations and cut points were developed and cross-validated.

Results: 196 subjects participated. One participant dropped out prior to testing due to time constraints, 15 CHD, 32 CF, 31 JDM, 31 JA, 30 IMD, 28 HE and 29 healthy controls completed the study. Agreement between predicted and measured EE varied across disease group and ranged (ICC) 0.13-0.46. Disease specific prediction equations exhibited a range of results (ICC 0.62 to ICC 0.88) (SE 0.45-0.78).

Conclusion: Poor agreement was demonstrated using current prediction equations in children with chronic conditions. Disease-specific equations and cut points were developed.

Acknowledgements: This study was funded with a grant from the Canadian Institute of Health Research (# 167391/CIHR). The study authors have no conflicts of interest to report. Manuscript accepted by Pediatric Exercise Science. Copyright Human Kinetics.

Introduction

Physical activity participation at moderate and vigorous levels has been shown to have an important protective effect on cardiovascular and metabolic profiles in healthy children that may prevent chronic illness later in adulthood (4, 17). The role of physical activity in children with chronic illness may have an even more substantial and wide-ranging health impact—from improving cardiometabolic profiles to improving physical function and limiting long-term disability (16). For example, accurate determination of physical activity related energy expenditure could allow for the better clinical management of diseases such as Cystic Fibrosis or Duchene's muscular dystrophy where nutritional considerations and energy requirements play a critical role in health outcome (36, 47). Further, the ability to accurately quantify physical activity dose (e.g. time spent in light, moderate and vigorous physical activity) in children with chronic conditions in order to prescribe a dose that confers benefit rather than harm is also of great importance. The true effect of physical activity level and role of physical activity related energy expenditure on health outcome in children with chronic conditions is currently unclear due, in part, to the lack of a precise measure of physical activity (16, 27, 40).

Objective methods to determine physical activity levels in children may provide more accurate and unbiased estimates compared to subjective measurement methods. One objective method that has gained popularity, accelerometry, has demonstrated superior validity and reliability when compared to subjective methods such as self-report (2, 38). Accelerometers have proven to be reliable and accurate in healthy children and adolescents when compared with doubly labelled water, calorimetry and direct observation (22, 25, 44). Raw signals related to change in velocity over time (m/s^2) are

collected over a designated period of time (epoch) and are converted into a unit-less index referred to as a count. Prediction equations are applied to convert counts into more physiologically meaningful indices of energy expenditure such as a metabolic equivalent (MET). Additionally, cut-point values enabling the relationship between counts and physical activity intensity to be established have also been developed (44). These equations and cut-point estimates have been validated for use in healthy children but not in those with chronic conditions.

Prediction equations and cut-point estimates developed to provide more meaningful indices of energy expenditure or physical activity level from accelerometer counts in healthy children may not be accurate in children with chronic conditions. Inaccuracies may arise from differences in resting metabolic rates, economy of motion, movement biomechanics, and energy cost of breathing in children with chronic conditions (35, 40). For example, a longitudinal study of changes in walking related energy expenditure in boys with Duchenne's reported increases in energy expenditure with a decrease in comfortable walking distance (8). Boys who were older in age (>8years) had the highest increases in walking energy expenditure in comparison to younger boys, a finding that is opposite of what occurs in typically developed children(8). Children with disease often have inherent physiological or biomechanical differences from healthy children (gait deficiencies or metabolic abnormalities) that require them to work at a higher energy level to complete the same task (16, 18, 22, 35). For example, a recent gait analysis study of 36 children diagnosed with polyarticular arthritis revealed significant differences in walking mechanics including: slower walking speeds, a shorter stride length, reduced hip and knee extension, tilted pelvis and improper foot mechanics when compared to healthy

controls(23). These changes may lead to a higher energy cost of locomotion; such differences may not be accounted for by the currently available prediction equations and cut-point estimates and could produce significant bias in determining energy expenditure or activity levels if used in their current form.

Objective tools to measure physical activity in children with chronic conditions are needed to accurately determine the relationship between energy expenditure, activity level and health outcomes in clinical populations. Therefore the aim of this study was to determine if the current prediction equations used to transform accelerometer counts into units of energy expenditure are valid in children with chronic conditions. A secondary aim of this study was to perform exploratory analyses in order to determine new prediction equations and cut-point estimates should the current equations not be accurate for use in children with a chronic disease.

Methods

A total of 167 subjects aged 7 to 18 years with a diagnosis of either: juvenile arthritis (JA), juvenile dermatomyositis (JDM), moderate or severe haemophilia (HE) A or B (up to 5% clotting factor), muscular dystrophinopathies (Beckers, Duchenes or other inherited muscle disease; IMD), cystic fibrosis (CF), or congenital heart disease (CHD) who had undergone a heart repair (Fontan repair or Tetralogy of Fallot repair) were approached to participate in the study with the approval from the research ethics boards representing three pediatric tertiary care hospitals located in North America and Europe. In addition, 29 healthy children, identified from study advertisements posted on internal hospital websites and newsletters, were recruited from the same centres. Children were

excluded if they were on medication that would affect the heart-rate response to exercise, such as beta-blockers, or if they were unable to ambulate or unable to cooperate with the testing procedures. Written informed consent and assent were collected from each guardian/subject in accordance to the guidelines of the ethics boards of the participating centers.

Exercise Testing

Each 2-hour testing session consisted of anthropometric measures, pulmonary function testing and three energy expenditure protocols: resting energy expenditure (REE), activities of daily living and exercise. Subjects were asked to abstain from eating or drinking anything but water for 2 hours prior to their appointment for more accurate determination of REE (30, 33, 34). In order to allow for the comparison of our study population with other clinical populations, each subject completed a number of questionnaires pertaining to their physical function and quality of life. All subjects underwent a familiarization session in which they were introduced to all equipment and were given instructions on how to complete each task.

Exercise Testing Procedures

At each session, height (Harpendum Stadiometer, London, UK) and weight (SR555 Stand-on Scale System, SR Instruments, Tonawanda, NY) were measured to the nearest 0.1 cm and 0.1 kg, respectively, with subjects wearing light clothing and no shoes. BMI was calculated as (kg/m^2) with BMI percentile determined for each subject (21, 28). Skinfold measures were taken at the triceps, biceps, subscapular and suprailiac locations with the sum of skinfolds (mm) determined. REE was measured for 20 minutes using

either a K4B2 (Cosmed, Albano Liziiale, Rome) or Cortex Metamax (Cortex Metamax, CORTEX Biophysik GmbH, Leipzig Germany) while the subject was resting quietly, but awake, in the supine position (30, 33, 34). Subjects then completed a series of 7 tasks to represent a broad spectrum of energy expenditure intensities ranging from those associated with activities of daily living to moderate and vigorous exercise. Tasks consisted of 6 minutes each of computer work, sweeping, over-ground walking at a self-selected pace, stair stepping and 3 treadmill walking tasks. In order to account for a wide variability in physical function level in our subjects, two different treadmill protocols were used. A low and high function treadmill walking protocol was each determined from the 6-minute self-selected walking task. Subjects who averaged a walking pace of ≥ 3.5 km/hr completed the high function treadmill protocol of 2.0 km/hr, 3.5 km/hr, 6.0 km/hr and those in the low function protocol completed a treadmill protocol of 1.0 km/hr, 2.0 km/hr, and 3.5 km/hr.

Outcomes

Indirect Calorimetry: Expired gases were collected during each task using a portable metabolic cart (K4b2 in North America or a Cortex Metamax in Europe). Prior to study initiation the inter-unit reliability was established and no correction factor was necessary given the high level of agreement, with an Intraclass Correlation Coefficient (ICC) of 0.98 (Confidence Interval 0.76-0.99) between the two metabolic systems. Oxygen uptake (VO_2) and carbon dioxide production (VCO_2) were recorded on a breath-by-breath basis and were averaged at 15-second intervals. Prior to each test the equipment was calibrated using standard protocols. Energy expenditure (J/kg/min) was calculated according to the de Weir equations (46). REE was calculated by taking the average of the last 10 minutes

of the values recorded by the metabolic cart during the 20 minute resting task. Activity Energy Expenditure (AEE) was calculated using the equation $AEE = Total\ Energy\ Expenditure\ (TEE) - measured\ REE$ (31, 32). Metabolic equivalents (TEE/REE) were also determined for each activity.

Accelerometry: Both the Actical (Mini Mitter, Co) and Actigraph 7164 (Actigraph, LLC) accelerometers were chosen to measure acceleration activity counts based on their use in national studies of health indices (e.g. National Health and Nutritional Examination Survey and Canadian Healthy Measures Survey). The Actical is an omni-directional monitor with a sampling frequency of 32 Hz and sensitivity of 0.01 g. Actigraph is a unidirectional monitor with a sampling frequency of 10 Hz and sensitivity of 0.05-2.0 g (10).

The two different accelerometer models, Actical and Actigraph, were placed on the right hip at the midaxillary line of the superior iliac crest using an elastic belt. The sampling rate was set to 15-second epochs for both devices. Prior to the initiation of this study, inter-monitor reliability was assessed and excellent agreement was established for all of the accelerometers distributed between the North American and European sites with ICC's ranging from 0.98-0.99 (Confidence Interval, 0.88-1.0). *Heart rate* measured beat by beat was determined by a heart rate monitor (Polar Vantage XL 650i, Polar Instruments, Kempele, Finland) for each task with data averaged each 15 seconds and downloaded for analysis.

Questionnaires:

Subjects completed the Child Health Assessment Questionnaire (CHAQ)(20, 37), the Quality of My Life Scale (QOML)(26), the Pediatric Quality of Life Scale (PedsQL)(45), and the Tanner Staging Questionnaire(15). The CHAQ consists of questions from 8 functional domains comprising eating, dressing and grooming, walking, arising, hygiene, reach, activities, and grip. A summary score is given based on the 8 functional activity domains and is rated on a 3-point scale (where 0 indicates no limitations and 3 indicates severe limitations). Overall quality of life (QOL) and health related quality of life (HRQOL) were measured by the QOML on separate 10 cm visual analog scale with lower scores indicating worse QOL/HRQOL and higher scores indicating better QOL/HRQOL. The PedsQL generic core, validated in childhood disease, consists of 23 items from 4 different dimensions related to physical, emotional, social and school functioning (45). Items are scaled on a 5-point ordinal scale with 0 indicating never and 4 indicating almost always. Items were reverse scored and transformed into a 100 point scale, a total score was derived by taking the sum of all of the items and dividing by the total number of completed items (out of 23) with higher scores indicating better quality of life(45). Pubertal stage was determined using a self-report Tanner staging method (15) . Pain and illness severity were determined using a 10 cm visual analog scale from the CHAQ questionnaire.

Data management

Data from the portable metabolic carts and from the 2 accelerometers were all imported into Microsoft excel and merged with the count data from the accelerometers for each task. A special programming script was written in Python (www.python.org) and run to

merge all of the subject accelerometer and metabolic data into one summary datasheet. In order to determine agreement between predicted and measured EE we averaged the accelerometer and metabolic data over the last 2 minutes of each of the task in order to capture steady state similar to the protocols of the original validation studies(22, 31, 32). For the purpose of the exploratory analysis, metabolic and accelerometer data were kept in an *as per monitor* format. The as per monitor data format was used for analysis purposes and kept the metabolic and accelerometer data in its' original 15 second format with the exception of subtracting out the last 30 seconds of each subjects' data to allow for transitions and timing issues between tasks. Data in this format were used to preserve data integrity and to encourage a more realistic picture of physical activity (6, 39). All individual items and scores for each of the questionnaires were entered into a Filemaker Pro database (FilemakerPro V., Filemaker, Inc., Santa Clara, Ca). Finally, all data were imported into SAS 9.3 (SAS Inc, Cary, N.C.) for further statistical analyses.

Statistical Analysis

Patient characteristics and questionnaire scoring were summarized using descriptive statistics (e.g. mean, range and standard deviations). Groups were compared using Analysis of Variance; subsequent pairwise comparisons were corrected for multiple testing using Tukey's HSD. The primary objective of this study was assessed by determining the criterion validity of estimated EE, calculated using manufacturer-specified accelerometer prediction equations (11, 22, 31, 32), and the reference standard of measured EE from indirect calorimetry using an Intra Class Correlation Coefficient (ICC) with 95% confidence interval calculated separately for the Actigraph and Actical. According to de Vet et. al, criterion validity is defined as, 'the degree to which scores of a

new measurement device agree with the scores of a gold standard' (13) and 'criterion validity may be assessed by intraclass correlation coefficients when the scores are expressed in the same units' (13). The level of agreement between scores of the instrument (in our case existing EE prediction equations) and the scores (EE) derived from the gold standard (indirect calorimetry) allows one to determine if the new instrument is a suitable substitute for the old (7). An ICC of 0.61–0.80 was selected a priori to indicate substantial agreement while an ICC >0.8 indicated excellent agreement (29). Equations that were used to calculate predicted energy expenditure included those of: Puyau et al.(31) for the Actical accelerometer and Corder et al. (11) , Freedson et al.(22) and Puyau et al. (32)for the Actigraph accelerometer. Evenson et al. (19) do not provide EE prediction equations in their calibration study, however, cut points from their study come with the strongest recommendation for use in healthy children (43). In order to assess classification accuracy (sensitivity and specificity), our disease specific cut points were compared to those of Evenson et al. (19).

Exploratory analyses to develop new prediction equations to estimate metabolic equivalents (METs) were conducted using mixed regression modelling (also referred to as Cross Sectional Times Series analysis) (9, 48) using SAS 9.3 (SAS Inc., Cary, N.C).

This method is better able to handle correlated data (the same outcome, energy expenditure, was measured multiple times across different tasks for each subject).

Accelerometer counts were log transformed to normalize their distribution.

Accelerometer counts of zero were modelled separately in our model such that a zero count was recognized as being equivalent to REE. Explanatory variables selected a priori were added to the models in a stepwise fashion to determine which contributed to the best

fit. Each model was evaluated using four model fitting approaches, comprising the Intra Class Correlation Coefficient (ICC), Akaike information criterion (AIC), Bayesian information criterion (BIC) and the standard error of the predicted value. A model with an ICC of >0.8 indicated excellent fit, and an ICC range of 0.61-0.80 indicated good fit. Limits of agreement were also calculated. Additionally, paired t-tests with Bonferroni correction for multiple testing were performed to ensure that EE determined from the new prediction equations were not different from the measured EE.

In order to determine how well our new prediction equations might predict energy expenditure outside of our sample we conducted a cross-validation using the leave one out technique(1). Mean absolute error and root mean squared error were calculated in order to evaluate the prediction error around our new equations.

Finally, cut points establishing the relationship between accelerometer counts and physical activity level, were determined using Receiver Operating curve (ROC) analysis with Prismlab 6.3 (Graphpad Software, Inc). Cut point thresholds that exhibited the best sensitivity and specificity for establishing sedentary, light, moderate and vigorous physical activity levels were determined.

Results

A total of 196 subjects were enrolled into the study; 1 subject dropped out prior to testing due to time constraints. Fifteen children with CHD, 32 CF, 31 JDM, 31 JIA, 30 IMD, 28 HE and 29 healthy subjects completed the study. Demographic information along with subjective ratings of physical function, pain, illness and quality of life for healthy and chronically ill subjects is presented in Table 1. Twenty-four of the 195 children

underwent the low-function group protocol: 1 CF, 4 JDM, 1 JIA, 4 HE, 13 IMD, and no CHD subjects. One healthy control also underwent the low function protocol as a result of a slow pacing during the 6-minute self-selected walking task. The CHAQ resulted in an overall ranking of mild disability for subjects with CF, HE, JA and JDM; mild to moderate disability for the IMD subjects; and no disability in subjects with CHD or healthy controls (14). Ratings of pain and illness were highest amongst the JA group. Subjects with CF reported that they had the worst overall health and IMD subjects reported the worst overall quality of life in comparison to other disease groups and healthy controls (See Table 1 for further results).

Those with a chronic disease had higher oxygen uptake across all of the tasks combined in comparison with healthy controls (ANOVA; $F_{(6, 55)} = 17.0, p < 0.0001$). The mean difference in oxygen uptake between the diseased subjects and healthy controls was 2 ml/kg/min with a range of 0.5 to 3 ml/kg/min. All disease classifications with the exception of the JDM group had energy expenditures that were significantly different from healthy controls. Mean VO_2 from the indirect calorimeter is plotted by physical task and by diagnosis in Figure 1.

Accelerometer counts differed amongst the disease groups as well as between the healthy controls with a mean difference across all tasks combined of 22 counts for both the Actical and Actigraph (range -8 to 89 counts). A significant difference in mean counts by diagnosis and task was found (ANOVA; $F_{(13, 55)} = 134, p < 0.0001$). Figure 2 illustrates the difference in mean accelerometer counts (Actigraph) by diagnosis across the physical tasks.

Agreement between the measured and predicted EE determined by an ICC and the 95% confidence interval around the ICC for Actical and Actigraph are presented in Table 2. Overall, when compared to measured, prediction equations for both the Actigraph and Actical resulted in poor to moderate agreement in the disease groups. The Puyau et al. equation (31) developed for the Actical resulted in better overall agreement than those for the Actigraph with the best performance in the CHD, HE and healthy groups. The Freedson equation (22) developed for the Actigraph 7124 had the best agreement across the disease groups in comparison to the other tested equations using the Actigraph.

Six disease specific equations were developed based on a significant interaction term between diagnosis and accelerometer count ($F_{(6,171)} = 4.3, p < 0.0005$) indicating that the relationship between counts and EE differed depending on diagnosis. All of the disease models included a random effect for the intercept and accelerometer count. Overall each of the disease specific models had an ICC value between 0.62-0.88 indicating good to excellent model fit. No significant differences between measured and predicted EE were found for any of the disease specific equations with the exception of the Actical equation that included heart-rate for IMD subjects (See results in Table 3).

Models that included heart rate in the prediction equation resulted in improved model fit and lower standard errors than prediction equations without heart rate. However, our disease specific prediction equations without heart rate still resulted in better agreement (ICC range 0.62-0.85) than the non-disease specific equations (ICC range 0.22-0.46) (see Supplemental Table 1 for disease specific ICC's and standard errors).

Results from our leave one out cross-validation suggest good prediction precision of our new equations. Generally prediction precision was similar across devices and disease groups with mean absolute errors ranging from 0.50 to 0.70; exceptions to this were CHD and JDM groups on using the Actigraph, which had poorer prediction precision (mean absolute error, 0.80 and 0.84, respectively). See Supplemental Table 2 for further cross-validation results.

Final regressions for Actigraph and Actical are presented both with heart-rate included and excluded by disease group (Table 3). Heart-rate was included in order to address clinical user groups who will use combined methods for determining energy expenditure. The prediction equations along with their limits of agreement are presented in Table 3.

ROC curve analysis results for sedentary, light, moderate, and vigorous intensity cut-points are presented in Table 4 for the Actigraph and Actical. The area under the curve ranged from 0.82 to 0.95 for Actical and 0.80 to 0.90 for Actigraph across the four intensity thresholds for all chronic diseases. Area under the curve and cut point results for individual disease groups are also presented in Table 4.

Evenson et al.(19) cut points were applied to our disease groups and resulted in similar specificity to our disease specific cut points. The sensitivity of the cut points were similar for sedentary activity but less than our disease specific cut points at higher levels of intensity. Further results are summarized in Supplemental Table 3.

Discussion

The findings of this study suggest that current prediction equations used to transform accelerometer counts into EE units are not accurate for children with chronic illnesses,

since the level of agreement was low to moderate (Table 2). Moreover, the agreement between predicted and measured EE was also only moderate in our healthy control group.

We found that EE and accelerometer counts differ across the different disease groups and in comparison to healthy controls (Figure 1 and 2). Differences in EE and accelerometer counts could be attributed to factors related to disease pathophysiology, sequelae, and drug treatment (41, 42). For example, mechanical efficiency due to joint damage or contractures or muscle atrophy from medication and muscle disuse in JA or JDM (23, 41); decreased lung function in CF (35) and disturbed muscle metabolism in IMD as well as deconditioning or poor neuromotor control may explain the differences in patterns of movement and increased energy cost of performing physical activities(8, 24). These differences likely explain the poor agreement using current prediction equations in our clinical populations. In a validation study of 34 adults with coronary artery disease poor agreement between predicted and measured EE using the Actigraph (7164) was also reported (18). The authors cited deconditioning and limits to functional capacity as one of the potential factors that contributed to the poor agreement of the prediction equations (18).

The current prediction equations resulted in a moderate level of agreement in subjects with HE and CHD. These groups may be more similar to healthy controls. Subjects with CHD in our study sample reported no disability and differences between our CHD and healthy controls may not have been evident due to the submaximal nature of our protocol(5). Recently, it has been reported that children who have undergone a Fontan procedure have an enhanced submaximal work capacity (lower EE) in comparison to their healthy peers, and that differences in oxygen uptake are only evident at a peak

workloads(5). The main sequelae of HE are related to the inability to produce clotting factor. The joint contractures and sequelae historically due to repeated significant bleeds have been largely eliminated in stable and well-managed North American and North Western European subjects with the advent of prophylactically replaced factor VIII and VIII.

A moderate level of agreement was also found using the current equations in our healthy controls. Our findings in the healthy group are in agreement with those published in the literature(3, 44). These studies also tested the validity of our selected published regression equations for the Actigraph and Actical and found that the accuracy of the equations was dependent on the type and intensity of the physical activity (3, 44). For example, Alhassan et al. found that the Freedson equation used for the Actigraph 7164 tended to overestimate high intensity activities in their sample of healthy children by as much as 26%, but was accurate for slow walking (3). The same group reported that the regression equation developed by Puyau et al, for the Actical underestimated EE across all activities (3, 44). There are several explanations for the moderate level of agreement found in our healthy subjects. A lack of substantial agreement could be due to differences in biological and physiological characteristics of our sample population or differences in protocols. It is possible that small validation samples in addition to subject selection pools may account for some of the lack of agreement in EE in our healthy control group. For example, a relatively small sample size was used to develop the original prediction equations and may make it difficult to capture the vast variability in EE in a different population (3). Biological characteristics of the study population such as age of the study participants may also account for some of the differences. Age ranges

were similar in our protocol to that of Freedson's group (age range 6 to 18) but differed from that of Puyau (age range 10 to 14)(3, 22, 44). Differences in the protocols used to develop the regression equations for the Actigraph also differed in our study as compared to Freedson's as activities of daily living were not included in the original Freedson validation protocol (22).

An exploratory analysis was conducted to address the need for more accurate EE prediction equations for children with chronic conditions. A new prediction equation was not developed for healthy children as much work has already been completed in this area with recent advances in regression modelling including time-series models, branched and 2 regression models demonstrating promising accuracy in this cohort (9, 12, 48) Our new disease specific models exhibited good to excellent model fit as assessed by ICC, low standard errors around the predicted value, and good precision accuracy determined by our cross-validation.

Our disease specific prediction equations were not significantly different from measured with one exception. A significant difference was found between measured and predicted EE for the IMD equation this finding may be by chance alone or could represent a true difference. The EE between predicted and measured was statistically different, however, this equation demonstrated good agreement, small mean differences, limits of agreement and standard errors, suggesting good clinical relevance.

In order to address the finding of a significantly different accelerometer counts across the physical tasks between disease conditions and also with healthy controls we developed new cut-point thresholds. ROC curve analysis was conducted to determine cut-point thresholds for light, moderate and vigorous physical activity for each disease group. The

overall area under the curve for all chronic diseases demonstrated moderate to excellent sensitivity and specificity. The classification accuracy of our equations is similar to findings using the same devices in the healthy population (3, 44). In their studies good classification was shown for identifying vigorous levels of physical activity with moderate to good classification accuracy for light activities and vigorous activities (3, 44).

The classification accuracy of our cut points performed better than those of Evenson et al., which are often recommended for determining intensity thresholds in healthy children(19). While the specificity of the Evenson cut points were similar to that of the disease specific cut points, the sensitivity (false-negative rate) was better in our disease specific cut points. Overall, this suggests that cut points recommended for healthy children would result in misclassification of moderate and vigorous physical activity as light or sedentary in children with a chronic condition.

Our results must be considered in the context of some potential limitations in our design. We looked at widely used prediction equations to derive EE from both the Actigraph and Actical accelerometer; it is possible that a different prediction equation might provide better agreement for children with chronic disease. However – it is unlikely that any prediction equation derived from healthy children would provide good accuracy; we have found a significant interaction between accelerometer counts and diagnosis that suggests that the relationship between accelerometer counts and EE is different in children with chronic conditions as compared to healthy controls and when compared to other disease groups (42). As an example, in our previous study we tested the validity of the Actiheart

accelerometer in the same population and found that a newly developed prediction equation tended to over-estimate energy expenditure in our chronic disease group(42).

We performed exploratory analyses to develop six disease specific prediction equations. While we used diagnosis as a means to break apart the prediction equations it is possible that diagnosis is actually representative of some other factor (so far unmeasured) that could also explain the differences found between these groups. Our findings suggest that disease specific energy expenditure equations and cut points perform better than those developed in healthy children. This could be because children with a chronic disease are actually different from healthy children or the equations and cut points work better because we have optimized them for our sample. While we encourage the use of these equations in preference to ones developed in the healthy population, further external cross validation work is required.

There are several strengths to our modelling approach that may overcome the limitations of other approaches. This type of modelling approach (sometimes referred to as cross-sectional time series analysis) (9, 48) has been used successfully in healthy cohorts and has shown great accuracy in predicting energy expenditure with reported mean percent errors ranging from ± 7.4 to 10.8 and correlation concordance coefficients ranging from 0.91 to 0.93 indicating near perfect agreement between measured and predicted energy expenditure(9, 48). The use of a mixed model using random coefficients better accounts for the correlation between EE present when a subject repeats physical tasks and may be used to better understand how well the modelled prediction will account for the variance in EE in the larger population (48). Also, we included all data points in our models rather than averaged data. Steady state data is often used to model prediction equations, but

may not properly represent what actually occurs in the field when determining EE. This approach is supported by the findings of others who have presented models that have improved performance when determining EE by using full signals(12, 48). Our leave one out cross-validation further supports the accuracy of the proposed prediction equations. Finally, we have provided cut-points for establishing physical activity level by disease group and accelerometer type that have shown good to excellent predictive capabilities.

In conclusion, our study has shown that current prediction equations used to determine EE from the Actical and Actigraph accelerometers are not valid in children with chronic conditions. We have completed preliminary work to develop six disease-specific prediction equations that have been shown to have good to excellent model fit and low standard errors around predicted METS. Further, we have shown good prediction accuracy of our prediction equations using a leave one out cross validation. We have also provided disease specific cut point values for the Actical and Actigraph for establishing physical activity levels. Future work should concentrate on confirming these equations and cut point values in other cohorts of children with chronic conditions and using them to establish relationships with health outcomes.

References

1. Cross Validation. 2008. In: SAS/STAT® 92 User's Guide [Internet]. Cary, NC: SAS Institute Inc.: SAS Institute Inc. Available from: [http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm - statug_glmselect_sect025.htm](http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm-statug_glmselect_sect025.htm).
2. Adamo KB, Prince SA, Tricco AC, Conner-Gorber S, Tremblay MS. A Comparison of indirect versus direct measures for assessing physical activity in the pediatric population: A systematic review. *International Journal of Pediatric Obesity*. 2009;4(2):2-27.
3. Alhassan S, Lyden K, Howe C, Keadle SK, Nwaokelemeh O, Freedson PS. Accuracy of Accelerometer Regression Models in Predicting Energy Expenditure and METs in Children and Youth. *Pediatric Exercise Science*. 2012;24:519-36.
4. Anderson LB, Harro M, Sardinha LB, Froberg K, Ekelund U, Brage S, et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *The Lancet*. July 22, 2006;368:299-304.
5. Banks L, McCrindle B, Russel JL, Longmuir PE. Enhanced physiology for Submaximal exercise in children after the Fontan procedure. *Med & Sci in Sports & Ex*. 2013;e published ahead of print.
6. Basset DR, Rowlands A, Trost SG. Calibration and Validation of Wearable Monitors. *Med & Sci in Sports & Ex*. 2012;44(1S):S32-S8.
7. Bland MaA, D. Statistical Methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986 February 8, 1986:307-10.
8. Brehm M-A, Kempen JCE, van der Kooi AJ, de Groot IJM, van den Bergen JC, Verschuuren JJGM, et al. Age-Related Longitudinal Changes in Metabolic Energy Expenditure during Walking in Boys with Duchenne Muscular Dystrophy. *PLoS ONE*. 2014;9(12):e115200.
9. Butte NF, Wong WW, Lee JS, Adolph AL, Puyau MR, Zakeri IF. Prediction of Energy Expenditure in Preschoolers. *Medicine and Science in Sport and Exercise*. 2014;46(6):1216-26.
10. Chen KY, Bassett, D.R. . The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. 2005;37(11):S490-S500.

11. Corder K, Brage, S., Wareham, N.J., Ekelund, U. . Comparison of PAEE from combined and separate heart rate and movement models in children. *Med Sci Sports Exerc.* 2005;37(10):1761-7.
12. Crouter SE, Erin E, J.D. H, Frongillo EA, Basset DR. A Refined 2-Regression Model for the Actigraph Accelerometer. *Med & Sci in Sports & Ex.* 2010;42(5):1029-37.
13. de Vet HCW, Terwee C, Mokkink LB, Knol D. Validity. 2011. In: *Measurement in Medicine [Online] Practical Guides to Biostatistics and Epidemiology [Internet].* Cambridge University Press; [150-201].
14. Dempster H, Porepa M, Young N, Feldman BM. The clinical meaning of functional outcome scores in children with juvenile arthritis. 44. 2001;44(8):1768-74.
15. Duke PM, Litt, I.F., and Gross, R.T. Adolescent's Self-Assessment of Sexual Maturation. *Pediatrics.* 1980 1980;66(6):918-20.
16. Durstine JL, Painter, P. Franklin, B.A. , Morgan, D. , Pitetti, K.H., Roberts, S.O. Physical activity for the chronically ill and disabled. *Sports Medicine.* 2000 2000 Sep;30(3):207-19.
17. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A. Moderate to Vigorous Physical Activity and Sedentary Time and Cardiometabolic Risk Factors in Children and Adolescents. *JAMA.* 2012;307(7):704-12.
18. Ekelund U, Tingstrom P, Kamwendo K, Krantz M, Nylander E, Sjostrom M. The Validity of the Computer Science and Applications activity monitor for use in coronary artery disease patients. *Clin Physiol & Func Im* (2002). 2002;22:248-53.
19. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *Journal of sport Sciences.* 2008;26(14):1557-65.
20. Feldman BM, Ayling-Campos A, Luy L, Stevens D, Silverman E, RM. L. Measuring Disability in Juvenile Dermatomyositis: Validity of the Childhood Health Assessment Questionnaire. *Journal of Rheumatology.* 1995;22(2):326-31.
21. Fredriks AM, Van Buuren S, Wit J, Verlooe-Vanhorick S. Body index measurements in 1996-7 compared with 1980. *Arch Dis Child.* 2000;82(2):107-12.

22. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc.* 2005;37(11 Suppl):S523-S30.
23. Hartmann M, Kreuzpointner F, Haefner R, Michels H, Schwirtz A, Haas JP. Effects of Juvenile Idiopathic Arthritis on Kinematics and Kinetics of the Lower Extremities Call for Consequences in Physical Activities Recommendations. *International Journal of Pediatrics.* 2010;2010.
24. Hausdorff JM, Zemaný L, Peng C-K, Goldberger AL. Maturation of gait dynamics: stride to stride variability and its temporal organization in children. *J Appl Physiol.* 1999;86:1040-7.
25. Hoos MB, Plasqui, G., Gerver, W.M. , Westerterp, K.R. . Physical activity level measured by doubly labeled water and accelerometry in children. *European journal of physiology.* 2003 2003 Apr-Jul;89:624-6.
26. Huber AM, Hicks JE, Lachenbruch PA, Perez MD, Zemel LS, Rennebohm RM, et al. Validation of the Childhood Health Assessment Questionnaire in the juvenile idiopathic myopathies. Juvenile Dermatomyositis Disease Activity Collaborative Study Group. *J Rheumatol.* 2001;28:1106-11.
27. Kesaniemi Y, Danforth E, Jensen M, Kopelman P, Lefebvre P, Reeder B. Dose-response Issues Concerning Physical Activity and Health: An Evidence-Based Symposium. *Med & Sci in Sports & Ex.* 2001;33(6 (Suppl)):S351-S8.
28. Kuczmarski R, Ogden C, Grummer-Strawn L, Flegal K, Guo S, Wei R, et al. *CDC Growth Charts: United States.* Hyattsville, MD: National Center for Health Statistics; 2000.
29. Landis JR, and Koch, G. The Measurement of Observer Agreement for Categorical Data. *Biometrics.* 1977 1977;33(1):159-74.
30. Pencharz PB, Azcúe MP. Measuring resting energy expenditure in clinical practise. *Journal of Pediatrics.* 1995;127(2):2135-43.
31. Puyau MR, Adolph, A.L. , Vohra, F.A. , Zakeri, I., Butte, N.F. Prediction of activity energy expenditure using accelerometers in children. *Med Sci Sports Exerc.* 2004 2004;36(9):1625-31.
32. Puyau MR, Adolph, A.L. , Vohra, F.A. , Zakeri, I., Butte, N.F. Validation and calibration of physical activity monitors in children. *Obesity Research.* Mar 2002;10(3):150-7.

33. Schmitz KH, Treuth, M. , Hannan, P., McMurray R., Ring, K.B., Catellier, D., Pate, R. . Predicting energy expenditure from accelerometry counts in adolescent girls. *Med Sci Sports Exerc.* 2005;37(1):155-61.
34. Schofield WN, Schofield, C., and W.P.T James. Basal Metabolic Rate: review and prediction, together with an annotated bibliography of source material. *Human Nutr Clin Nutr.* 1985;39C:1-96.
35. Shakur Y, Richards H, Pencharz P. Is it Necessary to Measure Resting Energy Expenditure in Clinical Practice in Children. *Journal of Pediatrics.* 2008;152:437-9.
36. Shimizu-Fujiwara M, Komaki H, Nakagawa E, Mori-Yoshimura M, Oya Y, Fujisaki T, et al. Decreased resting energy expenditure in patients with Duchenne muscular dystrophy. *Brain and Development.* 2012;34:206-2012.
37. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis Rheum.* 1994;37(12):1761-9.
38. Slootmaker S, Schuit A, Chinapaw M, Seidell J, Van Mechelen W. Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education, and weight status. *International Journal of Behavioral Nutrition and Physical Activity.* 2009;6(17).
39. Staudenmayer J, Zhu W, Catellier DJ. Statistical Considerations in the Analysis of Accelerometry-Based Activity Monitor Data. *Med & Sci in Sports & Ex.* 2012;44(1S):S61-7.
40. Strath SJ, Pfeiffer K, Whitt-Glover M. Accelerometer Use with Children, Older Adults, and Adults with Functional Limitations. *Med & Sci in Sports & Ex.* 2012;44(1S):S77-85.
41. Takken T, Elst E, Van der Net J. Pathophysiological Factors Which Determine Exercise Intolerance in Patients with Juvenile Dermatomyositis. *Current Rheumatology Reviews.* 2005;1:91-9.
42. Takken T, Stephens S, Balemans A, Tremblay MS, Esliger DW, Beyene J, et al. Validation of the Actiheart activity monitor for measurement of activity energy expenditure in children and adolescents with chronic disease. *Eur J Clin Nutr.* 2010;64(12):1494-500.
43. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of Accelerometer cut points for predicting activity intensity in youth. *Medicine and Science in Sport and Exercise.* 2011;43(7):1360-8.

44. Trost SG, Way, R. , Okely, A.D. Predictive validity of three ActiGraph energy expenditure equations for children. *Med Sci Sports Exerc.* 2006;38(2):380-7.
45. Varni J. The PedsQL Multidimensional fatigue scale in pediatric rheumatology: Reliability and Validity. *The Journal of Rheumatology.* 2004;31(12):2494-500.
46. Weir JBdV. New Methods for Calculating Metabolic Rate with Special Reference to Protein Metabolism. *J Physiol.* 1949;109:1-9.
47. Williams CA, Stevens D. Physical activity and exercise training in young people with cystic fibrosis: Current recommendations and evidence. *Journal of Sport and Health Sciences.* 2013;2:39-46.
48. Zakeri I, Adolph AL, Payau MR, Vohra FA, Butte NF. Application of cross-sectional time series modeling for the prediction of energy expenditure from heart rate and accelerometry. *J Appl Physiol.* 2008;104:1665-73.

Table 1: Descriptive characteristics of subjects, by disease category

Characteristic	CF	CHD	HE	IMD	JDM	JA	Healthy
Age, years							
Mean \pm SD	12.8 \pm 2.9	13.6 \pm 3.3	12.4 \pm 3.3	12 \pm 3.4	13.4 \pm 2.3	12.7 \pm 2.6	13.1 \pm 2.8
Gender, n (%)							
Male	18(56)	10(67)	28(100)	22(73)	11(35)	8(26)	16 (55)
Female	14(44)	5(33)	0 (0)	8(27)	20(65)	23(74)	13 (45)
Height, cm							
Mean \pm SD	156 \pm 16	161 \pm 17	156 \pm 19	146 \pm 22	159 \pm 11	154 \pm 12	162 \pm 16
Weight, kg							
Mean \pm SD	45 \pm 14	54 \pm 17	53 \pm 20.7	41 \pm 14	52 \pm 14	47 \pm 14	57.6 \pm 20
BMI percentile	3%	7%	4%	7%	10%	16%	0%
% Underweight <5%							
% Overweight 85-95%	0%	27%	29%	17%	13%	16%	28%
% Obese >95%	3%	7%	7%	7%	10%	3%	0%
Sum of Skinfold, mm							
Mean \pm SD	31 \pm 13	42 \pm 15.5	40 \pm 20	41 \pm 18	48 \pm 17	46 \pm 22	38 \pm 17
Tanner Stage (%)							
Prepubescent (Stages 1,2)	19%	38%	27%	70%	27%	32%	30%
Pubescent (Stages \geq 3)	81%	62%	73%	30%	73%	68%	70%
CHAQ Total Score	0.27 \pm 0.3	0.17 \pm 0.3	0.25 \pm 0.4	0.80 \pm 0.7	0.40 \pm 0.6	0.50 \pm 0.5	0.15 \pm 0.26
CHAQ Pain	0.6 \pm 0.9	0.2 \pm 0.6	1.0 \pm 1.7	1.5 \pm 1.8	1.4 \pm 2.1	2.6 \pm 3	1.3 \pm 1
CHAQ Overall Health	2.1 \pm 2.8	0.4 \pm 0.6	0.5 \pm 1.3	1.5 \pm 2	1.2 \pm 1.7	1.9 \pm 2.4	1.3 \pm 2
CHAQ Illness	1.8 \pm 1.9	0.6 \pm 1.3	0.7 \pm 1.4	2.0 \pm 2.4	2.0 \pm 2.1	2.1 \pm 2	1.3 \pm 2
PedsQL Total Score	78 \pm 12	72 \pm 12	82 \pm 16	68 \pm 17	77 \pm 15	72 \pm 13	83 \pm 9

CF = Cystic Fibrosis; CHD = Congenital Heart Disease; HE = Hemophilia; IMD = Inherited Muscle disease; JDM = Juvenile Dermatomyositis; JA = Juvenile Arthritis; CHAQ_DI = Child health Assessment Questionnaire; PedsQL= Pediatric Quality of Life Questionnaire; SD= standard deviation