Published in final edited form as:

Physiol Meas. 2016 December ; 37(12): 2231–2244. doi:10.1088/1361-6579/37/12/2231.

Filtering for productive activity changes outcomes in step-based monitoring among children

Michael Wininger^{1,2,3} and Kristie Bjornson⁴

¹Department of Rehabilitation Sciences, University of Hartford, 200 Bloomfield Avenue, West Hartford, CT 06117, USA

²Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA

³Cooperative Studies Program, Department of Veterans Affairs, West Haven, CT, USA

⁴Department of Pediatrics, Seattle Children's Research Institute, University of Washington, 2001 8th Ave, #629, Seattle, WA 98121, USA

Abstract

Wearable activity monitors are increasingly prevalent in health research, but there is as yet no data-driven study of artefact removal in datasets collected from typically developing children across childhood. Here, stride count data were collected via a commercially available activity monitor (StepWatch), which employs an internal filter for sub-threshold accelerations, but does not post-process supra-threshold activity data. We observed 428 typically developing children, ages 2–15, wearing the Step Watch for 5 consecutive days. We developed a minimum per-minute stride-count below which the data outputted from the StepWatch could be considered 'idle' and not 'productive'. We found that a threshold stride count of 10 steps per minute captured 90% of samples in a weighted average among isolated non-zero stride-count samples offset by inactivity. This threshold did not vary by age, gender, or by an age-gender interaction. Filtering the activity data according to this threshold reduced overall stride count by 8–10% by age group, from 8177 ± 2659 to 7432 ± 2641 strides per day. The impact on number of bouts per day decreased from an overall average of 79.3 ± 17.2 to 72.7 ± 12.1 ; this effect varied by age group. This study delivers the first data-driven estimate of a minimum activity threshold in step- or stride units that may extend to other studies. We conclude that the impact of production-idle filtering on activity data is substantial and suggests a possible impetus for re-contextualizing extant studies and guidelines reported without such filtering.

Keywords

activity; step; stride; pediatrics; activity monitoring; StepWatch; actigraph

Introduction

Activity monitoring is an increasingly prevalent and attractive paradigm for measuring health and well-being: wearable monitors are non-invasive, inexpensive, and allow data collection in native environments across many days (Patterson *et al* 1993, Tudor-Locke and Myers 2001, Taraldsen *et al* 2012). Studies incorporating stride-activity data, in particular, have yielded compelling insights into the behavior of actors of all ages and abilities (Cavanaugh *et al* 2007, Orendurff 2008, Adolph *et al* 2012, Danks *et al* 2014, Tulchin-Francis *et al* 2014), and provide invaluable data for the formulation of recommendations for minimum activity levels (Vincent and Pangrazi 2002, Tudor-Locke *et al* 2004, Graser *et al* 2011). However, where empirical activity data are becoming the basis on which science and policy is founded, it is critically important that these data be trustworthy and well-understood. Here we take a new look at conventional stride count data collected from an accelerometry-based device (StepWatch), and ask: should these data be filtered to account for putative artefact, and what is the impact of such filtering?

Typically among studies of large normative pediatric cohorts and youth with disability, activity data are not post-processed for a minimum activity level (Cavanaugh *et al* 2007, Dillon *et al* 2009, Bjornson *et al* 2010, 2014a, Sheiko *et al* 2012). Though less common, there is still some precedent for defining a minimum activity level (Orendurff 2008, Tulchin-Francis *et al* 2014), though the thresholds vary from study to study, see 1–2 min of ‘activity counts’ between up to 100 (Troiano *et al* 2008) and 500 per min (Tudor-Locke *et al* 2009). The plurality of thresholds may reflect their arbitrary nature, at best based on descriptive definitions of sedentary behavior (Matthew 2005); perhaps a data-driven approach would beget a more consistent rule. Furthermore, whereas these minimum activity count thresholds are given in units of ‘accelerometer counts’, they are much more likely to be device-specific, and are more obscure than a threshold based in units of step- or stride count. In this study, we seek to determine a threshold for the activity monitoring record in terms of a minimum stride-count; our goal is to propose an interpretable and scrutable minimum activity threshold for filtering activity records, in stride units, for application to typically developing children in generalized monitoring scenarios.

Our intention is to deliver a paradigm for filtering device data that may already be filtered. We presume that all devices filter data by magnitude: weak, sub-threshold signals are discarded autonomously by the device, and are never seen in the outputted data. However, to the knowledge of the authors, this filtering occurs instantaneously and sample-by-sample (or within a very short buffer comprising a few samples of data). We propose here to filter (or *re-filter*) data obtained from these devices, on a coarser scale, i.e. on the order of data reporting intervals (presumably at or near the scale of per-minute) rather than at or near the hardware sampling frequency (presumably tens or hundreds of samples per second). The filter proposed would identify and discard device data that may be reasonably assumed as not related to productive physical activity.

To accomplish this goal, we impose a new framework onto the activity log in order to categorize stride counts as either ‘productive’ or ‘idle’, i.e. as activity that either contributes or does not contribute to what would be considered purposeful or effortful action: isolated

data points with spuriously low stride-counts are common in the activity record, and would probably be considered non-productive activity by most clinicians (figure 1).

Despite their putative irrelevance to intentional action, they remain in the record, potentially biasing inferences drawn from activity monitoring datasets. In this study, we develop an empirical criterion for distinguishing these productive versus idle epochs, so that they can be filtered from the data record in a single processing step.

Methods

Subjects and protocol

This is a secondary data analysis of a cohort of 428 typically developing children between the ages of 2 and 15 years (Bjornson *et al* 2010). A minimum of 30 girls and 30 boys participated across seven age bands (2–3, 4–5, 6–7, 8–9, 10–11, 12–13 and 14–15 years). Recruitment and enrollment criteria have been previously published (Bjornson *et al* 2010). All participants wore the ankle mounted StepWatch on their left ankle for a minimum of seven consecutive days, all waking hours except when bathing or swimming. The StepWatch was individually calibrated per height and walking characteristics during a walking trial of >100 strides by a research assistant per published protocols (Bjornson *et al* 2007). Whereas at the time of data collection, there were neither manufacturer recommendations nor published protocols for data collection with the StepWatch in children, and given our design as cross-sectional across a range of ages and anthropometric sizes, all monitors were set to the following parameters: Quick Stepping = No, Walking Speed = Normal, Range of Walking Speeds = Moderate, and Leg Motions = Normal. Analysis was based on a 5 d sample (4 weekdays and 1 weekend day) of the seven day monitoring period (Ishikawa *et al* 2013).

Definitions and assumptions

The StepWatch activity monitor senses movement of the user relative to the environment, based on a continuous stream of digitized signal so from an embedded accelerometer. Within the StepWatch there is a ‘step determination unit’ which yields ‘counts’ of motions that breach a pre-defined threshold; this threshold is defined in gravimetric units (‘g-force’) and can be calibrated (see above: *Subjects and Protocol*). Any such supra-threshold acceleration is considered a step. Therefore, ‘Activity’ is defined as any data point for which at least one stride (foot leaves the ground to next contact with ground) is recorded. A ‘Bout’ is any period of activity offset by inactivity. In particular, we note that there is no on-board processing of activity count data within the StepWatch system: data outputted from the device reflects the number of supra-threshold accelerations per unit time without further filtering.

For the purposes of processing these data, we define a ‘Singleton’ as subset of bouts comprising only one data point. Our use of ‘singleton’ here differs from that used previously (Orendurff 2008), where it was taken to mean a single step—a singleton here can contain multiple steps. Furthermore, we define ‘Production’ as any activity that exceeds a threshold stride count. Stride counts below this threshold reflect a high likelihood of noise, artefact, or

activity that would not be considered veridical activity in most cases for healthy persons. Common sources of sub-threshold activity include: leg jerk due to sneeze, spasm, or reflex activation; change of body position at rest; and small short walks, e.g. from across a small room. For ease of reference, we shall refer to sub-threshold data points as 'Idle'. Idle and production are mutually distinct subsets of activity.

In order to estimate the threshold for idle versus production, we make the following assumption: the majority of singletons reflect idle activity; we also make a corollary assumption: singletons with larger sustained offsets are more likely to reflect idle than singletons with smaller offsets, i.e. a singleton preceded (and followed) by 10 min of inactivity is more likely to reflect idle than a singleton preceded (and followed) by 1 min of inactivity.

We exhibit these definitions in figure 2. Two bouts are shown (7:20pm–7:29pm, and 7:37pm–7:43pm), and two singletons are shown (7:32pm and 7:35pm). The earlier singleton is a 2 min singleton because it is offset by two minutes of inactivity both preceding and following the active minute; the latter singleton is a 1 min singleton: while two minutes of inactivity precede the active minute, only a single inactive minute follows; singleton order is defined by the lesser of the two offsets.

We define the production-idle threshold as the stride count criterion for classifying a data point as productive activity versus idle activity. In the next section, we describe our method for estimating this threshold.

Production-idle threshold

In the previous section, we declared the assumptions that (1) most singletons reflect idle activity, and (2) singletons with larger offsets are more likely to reflect idle activity than those with smaller offsets. Here, we set the parameter for Assumption 1 at 90%: a point-estimate for this threshold shall be given according to the stride-count at which 90% of all singletons have the same or fewer strides. In service to Assumption 2, the sample of singletons for estimating this threshold will be weighted in linear proportion to offset size. By way of example: in reference to figure 2, the 2 min singleton would be counted twice, and the 1 min singleton counted once.

Stride counts for all singletons for all days in a subject's record were so counted, and the 90th percentile stride-count value was retained as the threshold estimate for that subject.

Statistical analysis

Here, we test two hypotheses: (1) the threshold is non-zero in some subset of the subject pool, and (2) the threshold is sensitive to at least one of the physiologic variables collected here, i.e. Leg Length, BMI, Gender, and Age (as a categorical variable 'AgeGp' every 2 years: 2–3, 4–5, 6–7, 8–9, 10–11, 12–13, and 14–15 years). We tested these two hypotheses simultaneously using a general linear model (GLM) of the form

$$\text{Threshold} \sim \text{LegLength} + \text{BMI} + \text{AgeGP} \times \text{Gender}$$

where the link function specifying the relationship between random and systematic components was taken as Gaussian; weighted priors were not used and the fitting method was iteratively reweighted least squares. We purposely discarded race and socioeconomic status (SES) on the presumption of irrelevance; we discarded height and weight for redundancy with body mass index (BMI); we opted for AgeGP (a 7-level ordered variable) versus Age (a continuous variable), despite its implications for statistical power, for the sake of conformity to the preponderance of studies where subjects are so grouped. Where our subject pool spans all of adolescence including the ages where most typically developing boys and girls would enter puberty, and whereas girls mature earlier than boys, we posed age and gender as an interaction term.

Hypothesis 1—a non-zero threshold for productive versus idle activity—was tested first via simple paired *t*-test across the population: a comparison of number of steps observed in the filtered condition versus the number of steps observed in the unfiltered condition; variances were assumed to be homogenous, so the pooled variance was used. This *t*-test was one-sided under the assumption that filtering would reduce the number of steps reported. If the null hypothesis of threshold of 0 steps was not rejected, a more granular search would be performed in order to find whether there were any sub-groupings of participants for which the filtering was impactful. For Hypothesis 2, we assessed each term in the GLM for any $P < 0.05$, i.e. the significance on each term; any term with a significant *P*-value would provide basis for rejecting the null hypothesis that the threshold does not co-vary with any predictor variable.

Results

Basic descriptors

The participants recruited for this study are described in table 1.

As expected in this cohort, leg length was strongly associated with both height (correlation $\rho = 0.969$), and age ($\rho = 0.943$), and thus was eliminated as a co-variate from all analyses. Because BMI is a direct transformation of height and weight, BMI was used in place of height and weight as the preferred co-variate in all analyses. And because BMI was more-than-moderately associated with age ($\rho = 0.592$), it was discarded as a co-variate. We note that BMI trended with age in a piece-wise fashion, with a ‘break-point’ at 6 years.

Singleton analysis

The average number of singleton data points per day across all subjects was 26.9 ± 9.4 , and the average estimate of threshold was 9.0 ± 3.9 strides. Table 2 shows the average and standard deviation of these parameters by age group.

In an N-way analysis of variance (ANOVA) relating singletons per day to Age Group, we found that Age Group was significantly associated with singleton count ($P < 0.001$), as was the Age Group-By-Gender interaction; Gender as an independent effect was not significant. We tested for select pair-wise effects via *t*-test: significance of differences was assessed between adjacent age-groups within-gender (figure 3, left) and within age-groups between

gender (figure 3, right). In order to preserve small signals of potential interest, we report P -values uncorrected for multiple comparisons.

In general, we find that a consistent trend of increase in number of Singletons with age, with a non-incremental increase in both boys and girls between Age Group 3 and 4, and between Age Group 6 and 7; within age groups, only one Gender-difference was found: Age Group 6, where girls yielded significantly more Singletons than boys. Notwithstanding the significant association between Age Group and Singleton counts, an N-way ANOVA relating the threshold to Age Group and Gender, Age Group, Gender, and the Age Group-by-Gender interaction were not significant.

Thus, we chose a threshold of 10 steps as a uniform threshold across all age groups and in both male- and female subjects. All subsequent analyses reflect this value for the threshold.

Stride count

Overall, the average daily stride rate per subject decreased from 8177 ± 2659 to 7432 ± 2641 , difference of 745 ± 162 , $P < 0.001$ in a paired t -test. The greatest reduction in a single subject was 1422 strides per day (Age = 10 years, Female), and the single smallest reduction was 243 strides per day (Age = 15 years, Male).

We tested for differential filtering effects among sub-groups via an N-way ANOVA on stride count differences as a function of Age and Gender, as well as their interaction; none were significant. We show the age-group effect in figure 4. Prior to filtering, the stride count varies over age from 7978 ± 2204 (2–3 years old) to 6926 ± 2866 (14–15 years old), with a peak at 9486 ± 2815 strides per day (5–7 years old, figure 4, left); following filtering, the stride count varies from 7315 ± 2233 (2–3 years) to 6230 ± 2785 (14–15 years), with peak at 8693 ± 2824 strides per day (5–7 years, figure 4 center). The filtering effect ranges from 663 ± 127 (2–3 years) to 696 ± 181 (14–15 years), with a peak at 809 ± 147 strides per day (8–9 years, figure 4, right). π = production-idle threshold.

Bouts

Filtering for a minimum stride count can have the effect of either increasing or decreasing the number of bouts observed. Consider a hypothetical example, two 5 min bouts of activity: 13-3-13-3-13 and 9-9-9-9-9; filtering according to a threshold of 10 steps per minute will yield 13-0-13-0-13 and 0-0-0-0-0, i.e. one bout converts to three bouts, and one bout converts to zero bouts, respectively.

The average daily number of bouts per subject decreased from 79.3 ± 17.2 (unfiltered data) to 72.7 ± 12.1 (filtered), a difference of 6.7 ± 18.4 , $P < 0.001$ in a paired t -test. The greatest reduction in a single subject was 75.8 bouts per day (Age = 15 years, Male), and the single greatest increase was 40.8 bouts per day (Age = 7 years, Female).

While the effect of filtering can be to increase or decrease the total number of bouts, the average impact within each age group is near zero for age groups 2–3 years (-5.4 ± 9.3 bout decrease), 4–5 (-5.0 ± 11.3 bout decrease) and 6–7 (-4.8 ± 14.2 bout decrease), the average impact is to increase the number of bouts in all subsequent age groups: $+10.0 \pm 15.3$ (age 8–

9), $+13.4 \pm 17.9$ (age 10–11), $+18.1 \pm 17.4$ (age 12–13) and $+20.8 \pm 18.0$ (age 14–15, figure 5).

Discussion

Study validity

In this work, we pose a straight-forward question: ‘what is the threshold stride-count below which activity data is considered not informative in measuring the walking activities of typically developing children as they grow’. In answer to this question, we create simple definitions of productive versus idle behavior, and a construct a simple measure of the productivity-idle threshold. While *defining* is an inherently axiomatic process, and therefore a contentious enterprise, we believe we have created an intuitive, easily interpretable framework here that is scrutable, and readily avails to replication and independent inquiry.

The literature reports daily unfiltered pedometer determined step counts to range from 10 000 to 13 700 for boys and 8400 to 11 300 for girls (Tudor-Locke *et al* 2004, Cardon and De Bourdeaudhuij 2007, Hands and Parker 2008, Laurson *et al* 2008). In 2010, Tudor-Locke described walking activity by secondary analysis of a large population based sample of children and youth unfiltered and filtered (not counting steps during <500 activity count min^{-1}). This work documented an average reduction of walking activity counts with the Actigraph of 2600 steps/day for children (6–11 years) and youth (12–19 years) with an increase in sedentary time each day. Children averaged 12 500 steps/day with filtered Actigraph step data, in contrast to our filtered StepWatch data average of 16 354 steps/day (converted to steps via 1 stride = 2 steps). Youth averaged 10 000 per Actigraph filtered steps (approximately 20% reduction), while our sample averaged 14 864 steps per filtered StepWatch data (approximately 10% reduction). We attribute the differences in absolute step count to possible differences in the way that accelerometer counts are converted to stride- or step-count between the two different devices, and acknowledge the difficulty in comparing step/strides between studies using different devices. We attribute the differences in filtering impact to the difference in threshold, but again recognize that it is difficult to compare between studies given that the Actigraph was filtered by activity-count and not by stride- or step-count. We note also that the StepWatch is considered the most accurate pedometer on the market today (Foster *et al* 2005, Song *et al* 2006, Ainsworth 2008, Bassett and John 2010, Carr and Mahar 2012, Feito *et al* 2012) particularly with respect to walking across various speeds and wandering (Algase 2003, Karabulut *et al* 2005), and particularly in application to children in both free walking and running (Song *et al* 2006). While the ankle-worn StepWatch has been shown to be susceptible to error due to heel tapping and leg swinging in a ‘laboratory setting’ with subjects seated in a chair and moving to a metronome of 120 beats per minute (Karabulut *et al* 2005), however we believe this is not a particularly informative finding for three reasons: (1) the error was observed in adults whose limb dynamics and performance of non-ambulatory activities are most likely different from those of children, (2) the setting of high-frequency, timed toe-tapping with both heels making contact against the ground, or both legs swinging together, while seated, for three full minutes may serve well as a protocol for device testing, but may not be a behavior that occurs frequently in typically developing children, and (3) the investigators reporting the

error themselves claim that this is ‘unlikely to have a large impact’ on step counts over a 24 h period.

Indeed, the ankle-worn StepWatch is used as a ‘criterion pedometer’ against which other pedometers are (or should be) compared in development (Karabulut *et al* 2005, Bassett and John 2010, Schmidt *et al* 2011, Silcott *et al* 2011). The StepWatch was designed to capture ‘walking activity only’; the Actigraph was developed for ‘physical activity’ measurement, and only subsequent to market deployment were algorithms developed to count steps. Thus, the StepWatch has the highest suitability for this kind of study among commercially available monitors.

Preschool age children (4–5 years) walk approximately 8000 more steps/day as measured with the filtered StepWatch data of this paper than counts with the unfiltered Digiwalker SW-200 (Cardon and De Bourdeaudhuij 2007). Using filtered Actigraph data, Tudor-locke 2010 (Tudor-Locke *et al* 2009) showed peak walking levels among 6 year-olds, which is consistent with both our unfiltered (Bjornson *et al* 2010) and filtered StepWatch data presented here. The contrast in stride count sensitivity may be a reflection of different event detection paradigms, which may have different axes of measurement or calibration protocols. In summary, we feel that our data acquisition is validated in showing congruency with previous studies to the extent possible, as small differences in scale are easily explained. Furthermore, our approach is simple, intuitive, and emphasizes a data-driven threshold discovery.

Study novelty

Though we are not the first to address the question of filtering activity monitors, we are the first—to our knowledge—to propose a threshold that is (1) explicitly data-driven, (2) in units of step or stride per minute, and (3) looks critically at whether the threshold changes across a large and diverse cohort of typically developing children in such a wide age range.

In this work, we build on the premise that spontaneous, isolated instances of low step- or stride counts in the activity record as artefact and a potential source for bias in analysis of walking activity among typically developing children. We propose an effective and easily implemented signal processing step for filtering data in order to remove sub-threshold non-productive activity. Furthermore, we measure idle within a large cohort of children in a wide range of ages, yielding a broadly informative view into its prevalence and volume in a maximally generalizable study.

Relevance

Typically developing children and adolescents demonstrate physical activity patterns that are markedly different from adults (Armstrong *et al* 1990, Bailey *et al* 1995, Eisenmann and Wickel 2005, Bjornson *et al* 2010, 2014b, Tudor-Locke *et al* 2012); there is as yet no validation study to support use of filters developed in the study of adults in application to pediatric activity data.

All commercially available activity monitors require some filtering for signal artefact (Bussmann *et al* 2001). But the algorithms supplied by the manufacturer are embedded

within the device, and are proprietary: their impact is thus capricious and inaccessible. In particular, the filtering used in the StepWatch is based on thresholding the accelerometer signal, but other systems may use different settings or different approaches altogether. Demonstrably, they are insufficient to remove all noise: given the extensive literature base of inquiry into proper filtering of device data, and our own observations of unlikely data points from even very reliable devices (figure 1, where, for example, an adolescent participant's data purports that a single step was taken at 2:25am and at 2:26am), there is need to 'filter the filter'. Furthermore, the primary development work on most ambulatory devices, including the StepWatch (Bjornson *et al* 2010), was executed on adults and not children. Given that there is an expectation that the data obtained from these wearable monitors will need to be filtered post facto, we assert that evidence-based filter design is critically important to advancement of the field of activity monitoring.

Implications

While many studies have been published on walking activity monitoring in children, the scope of these studies is wide, and the protocols and devices are diverse and non-homogenous. In particular, activity studies also tend to report results in whatever is the base unit of measurement, which is often 'accelerometer count' and not step- or stride count. Thus, among those studies that do implement a minimum-activity filter, the filtering is also specified in the native unit of the device, which is not always in step- or stride units. As such, it is difficult to draw inferences about how our study may impact extant studies. However, we do propose that our method may help bring new life to previously published works: revisiting the conclusions drawn from unfiltered datasets may provide interesting new insights, or reinforce the existing conclusions. This opportunity extends to the interpretations of commonly used measures, specifically bouts: Whereas it is demonstrated here that the number of bouts can actually increase with filtering, we recognize that it is increasingly important to consider whether the concept of bout is informative without adjuvant context, e.g. bouts of activity lasting for more than a defined duration. We discuss this further (see below: *Limitations*). For ongoing and future studies, we urge that investigators consider implementing a productivity filter, or—if such a filter is not suitable—to at least acknowledge that filter was considered and to explain the reason for non-use.

In terms of policy-making, we observe the centrality of stride-count to guideline activity levels: A minimum of 60 min per day of moderate intensity physical activity has been recommended by the U.S. Surgeon general for optimal body mass (President's Challenge 2008). Elsewhere, it has been suggested that walking rates >120 steps min^{-1} can be considered moderate to vigorous physical activity (MPVA) (Graser *et al* 2011). Based on a cohort of 711 children, 6–12 years, Vincent and Pangrazi recommend 11 000 steps/day for optimal BMI (Vincent and Pangrazi 2002). Employing BMI-referenced data, Tudor-Locke recommended 12 000 steps/day for girls and 15 000 steps for boys based upon a pedometer determined international sample (Tudor-Locke *et al* 2004). Whether—and to what extent—this kind of filtering might provide evidence in support of guideline revision deserves further inquiry.

Limitations

Because of the utter novelty of our approach, and our focus on pediatric data, we could not find a literature base on which to draw immediate comparisons of our methodology. Likewise, because of the diffuse nature of the existing literature base, it is difficult to draw connections to other activity monitoring studies. For example, while activity monitoring in young athletes is compelling and interesting research, the experimental protocol and patient pool is simply too different to make meaningful comparisons to typically developing children or populations with disability (Orendurff *et al* 2010). And whereas we define 'bouts' in the most general way (any episode of activity offset before- and after by inactivity), many references to the concept of a bout are with a pre-defined minimum duration, e.g. 20 min (Masse *et al* 2005); this may complicate interpretation of our results to some previous works, but we make note that we were interested to draw comparisons to a particular set of papers where the reporting included bouts of any duration (Orendurff 2008, Bassett and John 2010, Tulchin-Francis *et al* 2014), a reflection of the fact that filtering could well change the duration of bouts in our data. Furthermore, we observe that our algorithm is intended to have impact in application of activity monitoring through step/stride count in conventional wearable accelerometers. Whether this approach would be augmented (or, perhaps: obviated altogether) by a device with the power to discriminate between body postures, e.g. activPAL (PAL technologies, Glasgow, UK), or by measurement of energy expenditure, remains an open and interesting question. Lastly, we note that this study is not designed to inform the classification of activity levels in the way of conventional low-, moderate-, and high activity; this study is intended to provide guidance on a post-processing filter for empirical artefact.

Future work

Real time videography of walking activity with coding for resulting levels of productive or idle walking would provide gold standard validation of this filtering approach to StepWatch data. Future validation work should also examine the relationship of the filtered walking activity data to other metrics of community-based mobility and participation. Previous work has shown that walking activity performance (unfiltered StepWatch data) is significantly associated with levels of participation in mobility based life habits for ambulatory children with cerebral palsy (Bjornson *et al* 2014b). Would those relationships be refined using filtered data? Similarly, the implementation of filtering in the processing of StepWatch data may influence the relative relationship of low, moderate and high walking activity by decreasing the number minutes in low activity (Bjornson *et al* 2014b). Such knowledge will inform the goal setting for populations with limitation in walking (i.e. obesity, cerebral palsy, spina bifida) as well as potentially direct interventions to optimize walking activity both for general health benefits and for guided rehabilitation.

Acknowledgements

Funding and support was provided by the Staheli Endowment Fund, Clinical Steering Committee Research Award, Department of Orthopedic Surgery, Seattle Children's Hospital, University of Washington School of Medicine-Medical Student Research Training Program, Seattle, Washington and CTSA 1 UL1 RR025014-01 (NCRR). The authors would like to acknowledge the support of the Research Summit IV: Innovations in Technology for Children with Brain Insults: Maximizing Outcomes, 15–17 October 2015, Alexandria, VA, Academy of Pediatric Therapy, American Physical Therapy Association.

References

- Adolph KE et al. 2012 How do you learn to walk? Thousands of steps and dozens of falls per day Psychol. Sci 23 1387–94 [PubMed: 23085640]
- Ainsworth BE 2008 How do I measure physical activity in my patients? Questionnaires and objective methods Br. J. Sports Med 43 6–9 [PubMed: 18718977]
- Algase DL 2003 Biomechanical activity devices to index wandering behaviour in dementia Am. J. Alzheimer's Dis. Other Dementias 18 85–92
- Armstrong N, Balding J, Gentle P and Kirby B 1990 Patterns of physical activity among 11 to 16 year old British children BMJ (Clin. Res. Ed.) 301 203–5
- Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ and Cooper DM 1995 The level and tempo of children's physical activities: an observational study Med. Sci. Sports Exercise 27 1033–41
- Bassett DR and John D 2010 Use of pedometers and accelerometers in clinical populations: validity and reliability issues Phys. Ther. Rev 15 135–42
- Bjornson K, Song K, Lisle J, Robinson S, Killien E, Barrett T and Zhou C 2010 Measurement of walking activity throughout childhood: influence of leg length Pediatric Exercise Sci. 22 581–95
- Bjornson KF, Belza B, Kartin D, Logsdon R and McLaughlin JF 2007 Ambulatory physical activity performance in youth with cerebral palsy and youth who are developing typically Phys. Ther 87 248–57 [PubMed: 17244693]
- Bjornson KF, Zhou C, Stevenson R, Christakis D and Song K 2014a Walking activity patterns in youth with cerebral palsy and youth developing typically Disab. Rehabil 36 1279–84
- Bjornson KF, Zhou C, Stevenson RD and Christakis D 2014b Relation of stride activity and participation in mobility-based life habits among children with cerebral palsy Arch. Phys. Med. Rehabil 95 360–8 [PubMed: 24231402]
- Bussmann JBJ, Martens WLJ, Tulen JHM, Schasfoort FC, van den Berg-Emons HJG and Stam HJ 2001 Measuring daily behavior using ambulatory accelerometry: the activity monitor Behav. Res. Methods Instrum. Comput 33 349–56 [PubMed: 11591066]
- Cardon G and De Bourdeaudhuij I 2007 Comparison of pedometer and accelerometer measures of physical activity in preschool children Pediatric Exercise Sci. 19 205–14
- Carr LJ and Mahar MT 2012 Accuracy of intensity and inclinometer output of three activity monitors for identification of sedentary behavior and light-intensity activity J. Obesity 2012 1–9
- Cavanaugh JT, Coleman KL, Gaines JM, Laing L and Morey MC 2007 Using step activity monitoring to characterize ambulatory activity in community-dwelling older adults: step activity monitoring in older adults J. Am. Geriatrics Soc 55 120–4
- Danks KA, Roos MA, McCoy D and Reisman DS 2014 A step activity monitoring program improves real world walking activity post stroke Disab. Rehabil 36 2233–6
- Dillon ER, Bjornson KF, Jaffe KM, Hall JG and Song K 2009 Ambulatory activity in youth with arthrogryposis: a cohort study J. Pediatric Orthopaedics 29 214–7
- Eisenmann JC and Wickel EE 2005 Moving on land: an explanation of pedometer counts in children Eur. J. Appl. Physiol 93 440–6 [PubMed: 15503123]
- Feito Y, Bassett DR and Thompson DL 2012 Evaluation of activity monitors in controlled and free-living environments Med. Sci. Sports Exercise 44 733–41
- Foster RC et al. 2005 Precision and accuracy of an ankle-worn accelerometer-based pedometer in step counting and energy expenditure Preventive Med. 41 778–83
- Graser SV, Groves A, Prusak KA and Pennington TR 2011 Pedometer steps-per-minute, moderate intensity, and individual differences in 12- to 14-year-old youth J. Phys. Act. Health 8 272–8 [PubMed: 21415454]
- Hands B and Parker H 2008 Pedometer-determined physical activity, BMI, and waist girth in 7- to 16-year-old children and adolescents J. Phys. Act. Health 5 S153–65 [PubMed: 18364519]
- Ishikawa S, Kang M, Bjornson KF and Song K 2013 Reliably measuring ambulatory activity levels of children and adolescents with cerebral palsy Arch. Phys. Med. Rehabil 94 132–7 [PubMed: 22892322]

- Karabulut M, Crouter SE and Bassett DR 2005 Comparison of two waist-mounted and two ankle-mounted electronic pedometers Eur. J. Appl. Physiol 95 335–43 [PubMed: 16132120]
- Laurson KR, Eisenmann JC, Welk GJ, Wickel EE, Gentile DA and Walsh DA 2008 Evaluation of youth pedometer-determined physical activity guidelines using receiver operator characteristic curves Preventive Med. 46 419–24
- Masse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ and Treuth M 2005 Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables Med. Sci. Sports Exercise 37 S544–54
- Matthew CE 2005 Calibration of accelerometer output for adults Med. Sci. Sports Exercise 37 S512–22
- Orendurff MS 2008 How humans walk: bout duration, steps per bout, and rest duration J. Rehabil. Res. Dev 45 1077–90 [PubMed: 19165696]
- Orendurff MS, Walker JD, Jovanovic M, Tulchin KL, Levy M and Hoffmann DK 2010 Intensity and duration of intermittent exercise and recovery during a soccer match J. Strength Cond. Res./Nat'l Strength Cond. Assoc 24 2683–92
- Patterson SM, Krantz DS, Montgomery LC, Deuster PA, Hedges SM and Nebel LE 1993 Automated physical activity monitoring: validation and comparison with physiological and self-report measures Psychophysiology 30 296–305 [PubMed: 8497559]
- President's Challenge 2008 Get Fit and Be Active! A Handbook for Youths Ages 6–17 (The President's Challenge) (Washington, DC The President's Council on Physical Fitness and Sports, US Department of Health and Human Services) (retrieved from www.presidentschallenge.org/tools-resources/docs/getfit_old.pdf)
- Schmidt AL, Pennypacker ML, Thrush AH, Leiper CI and Craik RL 2011 Validity of the StepWatch step Activity Monitor: preliminary findings for use in persons with parkinson disease and multiple sclerosis J. Geriatric Phys. Ther 34 41–5
- Sheiko M, Bjornson K, Lisle J, Song K, Eary JF and Conrad EU 2012 Physical activity assessment in adolescents with limb salvage J. Pediatrics 161 1138–41
- Silcott NA, Bassett DR, Thompson DL, Fitzhugh EC and Steeves JA 2011 Evaluation of the Omron HJ-720ITC pedometer under free-living conditions Med. Sci. Sports Exercise 43 1791–7
- Song KM, Bjornson KF, Cappello T and Coleman K 2006 Use of the StepWatch Activity Monitor for characterization of normal activity levels of children J. Pediatric Orthopaedics 26 245–9
- Taraldsen K, Chastin SFM, Riphagen II, Vereijken B and Helbostad JL 2012 Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: a systematic literature review of current knowledge and applications Maturitas 71 13–9 [PubMed: 22134002]
- Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T and McDowell M 2008 Physical activity in the United States measured by accelerometer Med. Sci. Sports Exercise 40 181–8
- Tudor-Locke C, Brashear MM, Katzmarzyk PT and Johnson WD 2012 Peak stepping cadence in free-living adults: 2005–2006 NHANES J. Phys. Act. Health 9 1125–9 [PubMed: 22207030]
- Tudor-Locke C, Johnson WD and Katzmarzyk PT 2009 Accelerometer-determined steps per day in US adults Med. Sci. Sports Exercise 41 1384–91
- Tudor-Locke C et al. 2004 BMI-referenced standards for recommended pedometer-determined steps/day in children Preventive Med. 38 857–64
- Tudor-Locke CE and Myers AM 2001 Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity Res. Q. Exercise Sport 72 1–12
- Tulchin-Francis K, Stevens W and Jeans KA 2014 Intensity and duration of activity bouts decreases in healthy children between 7 and 13 years of age: a new, higher resolution method to analyze StepWatch Activity Monitor data Physiol. Meas 35 2239–54 [PubMed: 25340461]
- Vincent SD and Pangrazi RP 2002 An examination of the activity patterns of elementary school children Pediatric Exercise Sci. 14 432–41

Figure 1.

Single-day exemplar of step count data shows subject most likely at rest between midnight and 6am, and again between 2pm and 3pm; total number of bouts = 116 (left). During presumptive windows of rest, subject accrues 22 bouts of five 'steps' or fewer, accounting for 19% of the day's bouts (right). Bout is defined as any period of activity offset by at least one minute (before and after) of inactivity.

Figure 2.
Exemplar data showing bouts and singletons.

Figure 3.

Two views into the interaction effects of Age Group and Gender in the number of Singletons per day: Grouped by Gender (left) and grouped by Age Group (right). * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$; Age Group 1 = 2–3 years old, Group 7 = 14–15 years.

Figure 4.

The number of strides by age group decreases by 8–10% per age group following filtering at a threshold of $\pi = 10$ strides.

Figure 5.

The impact of filtering on the number of bouts varies from 15.5% decrease to 8.5% increase.

Table 1.Basic demographic variables for participant pool ($n = 428$).

	Group size	Sex (M/F)	Leg length (in)	BMI
All participants	$n = 428$	216/212	28.6 ± 6.6	18 ± 3.3
Group 1 (2–3 years)	$n = 60$	30/30	18.4 ± 1.9	16.5 ± 1.3
Group 2 (4–5 years)	$n = 62$	31/31	22.5 ± 1.7	16.1 ± 1.5
Group 3 (6–7 years)	$n = 62$	32/30	26.5 ± 2.8	16.4 ± 2.1
Group 4 (8–9 years)	$n = 63$	32/31	28.9 ± 2.0	17.1 ± 2.5
Group 5 (10–11 years)	$n = 61$	31/30	31.9 ± 2.4	18.7 ± 3.0
Group 6 (12–13 years)	$n = 60$	30/30	34.9 ± 2.0	20.8 ± 3.6
Group 7 (14–15 years)	$n = 60$	30/30	36.9 ± 2.7	21.7 ± 3.4

Table 2.

Average number of singletons per day, and estimates of productive-idle threshold within age group; The threshold is defined as the number of strides of the 90th-percentile singleton.

	Singletons per day	Threshold estimate
All participants	26.9 ± 9.4	9.0 ± 3.9
Group 1 (2–3 years)	19.5 ± 4.9	9.8 ± 4.7
Group 2 (4–5 years)	21.5 ± 5.5	10.1 ± 4.3
Group 3 (6–7 years)	21.9 ± 6.3	8.8 ± 3.8
Group 4 (8–9 years)	28.2 ± 7.4	8.7 ± 3.2
Group 5 (10–11 years)	30.8 ± 8.8	8.6 ± 3.6
Group 6 (12–13 years)	32.8 ± 9.6	8.3 ± 3.5
Group 7 (14–15 years)	33.4 ± 9.9	8.4 ± 3.7