Medicine & Science IN Sports & Exercise
The Official Journal of the American College of Sports Medicine
www.acsm-msse.org

. . . Published ahead of Print

Comparability of Measured Acceleration from Accelerometry-Based Activity Monitors

Alex V. Rowlands[1,2], François Fraysse[2], Mike Catt[3], Victoria H. Stiles[4], Rebecca M. Stanley[2], Roger G. Eston[2,4], and Tim S. Olds[1]

[1]Health and Use of Time (HUT) Group, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia; [2]Exercise for Health and Human Performance Group, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia; [3]Institute for Ageing and Health, Newcastle University, United Kingdom;[4]Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, United Kingdom

Accepted for Publication: 14 May 2014

ACCEPTED

Medicine & Science in Sports & Exercise® **Published ahead of Print** contains articles in unedited manuscript form that have been peer reviewed and accepted for publication. This manuscript will undergo copyediting, page composition, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered that could affect the content.

Copyright © 2014 American College of Sports Medicine

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Medicine & Science in Sports & Exercise, Publish Ahead of Print
DOI: 10.1249/MSS.0000000000000394

Comparability of Measured Acceleration from Accelerometry-Based Activity Monitors

Alex V. Rowlands[1,2], François Fraysse[2], Mike Catt[3], Victoria H. Stiles[4], Rebecca M. Stanley[2], Roger G. Eston[2,4], and Tim S. Olds[1]

[1]Health and Use of Time (HUT) Group, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia; [2]Exercise for Health and Human Performance Group, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia; [3]Institute for Ageing and Health, Newcastle University, United Kingdom; [4]Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, United Kingdom

Conflicts of Interest and Source of Funding: Adult data (part 1) were part of a larger study funded by an Innovative Award from the National Osteoporosis Society, UK. There are no conflicts of interest. No external source of funding was received for this study.

Corresponding author:
Dr Alex Rowlands
School of Health Sciences,
University of South Australia
GPO Box 2471
Adelaide, SA 5001
Australia
Tel: +61 (0) 8830 21207
Email: Alex.Rowlands@unisa.edu.au

Running Head: Comparability of GENEActiv & GT3X+ output

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Abstract

Background: Accelerometers that provide triaxial measured acceleration data are now available. However, equivalence of output between brands cannot be assumed and testing is necessary to determine whether features of the acceleration signal are interchangeable. **Purpose:** To establish the equivalence of output between two brands of monitor in a laboratory and in a free-living environment. **Methods:** Part 1: Thirty-eight adults performed nine laboratory-based activities while wearing an ActiGraph GT3X+ and GENEActiv at the hip. Part 2: Fifty-eight 10-12 y old children wore a GT3X+ and GENEActiv at the hip for seven days in a free-living setting. **Results:** Part 1: The magnitude of time-domain features from the GENEActiv was greater than from the GT3X+. However, frequency domain features compared well, with perfect agreement of the dominant frequency for 97-100% of participants for most activities. Part 2: Mean daily acceleration measured by the two brands was correlated ($r=0.93$, $p<0.001$, respectively), but the magnitude was ~15% lower for the GT3X+ than the GENEActiv at the hip. **Conclusion:** Frequency-domain-based classification algorithms should be transferable between monitors and it should be possible to apply time-domain-based classification algorithms developed for one device to the other, by applying an affine conversion on the measured acceleration values. The strong relationship between accelerations measured by the two brands suggests habitual activity level and activity patterns assessed by the GENE and GT3X+ may compare well if analysed appropriately.

Keywords: GENEActiv; ActiGraph; GT3X+; time-domain; frequency-domain

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Introduction

The output from most accelerometry-based activity monitors (e.g. the ActiGraph GT1M, RT3, Actical) is reported in proprietary 'counts'. These counts are not comparable between different models of monitor due to differences in the processing, filtering and scaling of the measured acceleration signal (15). This has limited the comparability of data obtained from the various models of accelerometry-based activity monitor (7).

Over the past few years accelerometry-based activity monitors that allow access to the measured triaxial acceleration data have become commercially available, e.g. the ActiGraph GT3X+ and the GENEActiv. Access to the measured acceleration data should facilitate comparisons between outputs regardless of which of these monitors is used (10). However, as stressed by Welk et al. (15), equivalence of the measured acceleration output cannot be assumed and rigorous equivalency testing is necessary to determine whether, and under which conditions, outputs from these monitors are interchangeable. Various signal processing techniques are increasingly being used in an attempt to identify activity type from accelerometer data (e.g. 3, 8, 9, 11, 15, 17). Pattern recognition approaches use multiple frequency and/or time domain features from the total acceleration signal and from individual axes of acceleration. Frequency-domain features, such as Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT), are very commonly used in activity monitoring (see e.g. 16, 17) and have been shown to reliably discriminate between reasonably cyclic activities such as running and walking (16, 17). However, frequency domain features do not distinguish very well between activities having the same typical frequency but different intensity levels, e.g. fast walking and jumping. For these activities, time domain features are more discriminative (6). For pattern recognition approaches to be applicable to accelerations derived from more than one device, it is necessary to explore

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

which of the time and frequency domain features are equivalent between the monitors. At present, only preliminary data from small samples (≤10 participants) are available; these data suggest that acceleration magnitudes are lower for the ActiGraph GT3X+ relative to the GENEActiv (6, 10), but that features from the frequency domain compare well (6). A difference in acceleration magnitudes may impact on estimates of moderate-to-vigorous intensity activity from cut-points, suggesting cut-points may need to be brand specific.

The aims of this study were to establish the comparability of output from two widely used commercially available triaxial accelerometry-based activity monitors, the ActiGraph GT3X+ and the GENEActiv during controlled laboratory-based activities (part 1) and free-living (part 2). We hpothesized that: 1) the magnitude of accelerations at the hip measured by the ActiGraph GT3X+ would be lower than those measured by the GENEActiv for specific activities (part 1); 2) features from the frequency domain of the acceleration signal would not differ between the ActiGraph GT3X+ and the GENEActiv for specific activities (part 1); and 3) the magnitude of accelerations measured by the ActiGraph GT3X+ would be correlated with, but lower than the GENEActiv for free-living, (part 2).

Methods

Part 1: Laboratory-based

Part 1 investigated the comparability of features from the GT3X+ and the GENEActiv worn at the hip in adults in a controlled laboratory setting. The devices were compared using data as 'raw' as possible; that is, performing as little processing as possible on the data extracted

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

from the devices. The data were taken from a larger study investigating the use of accelerometry to classify activity beneficial to bone in premenopausal women (12).

Thirty-eight women (age: 39.3 ± 5.7 y, mass: 65.8 ± 11.8 kg, height: 1.67 ± 0.06 m) were recruited from the South West of the UK. Participants filled out a Physical Activity Readiness Questionnaire to confirm their ability to participate in the study. The institutional ethics committee granted approval and all participants gave written informed consent.

Data Collection: Each participant performed nine activities (slow walking, slow walking while carrying a weighted bag (2 kg) in the right hand, floor sweeping, brisk walking, slow running, faster running, low vertical jumps, higher jumps and box drops). For walking, sweeping and running activities, participants performed ten shuttles of 12-15 m each in length. The first shuttle of each ambulatory activity was performed at a self-selected cadence and a metronome was used to ensure subsequent shuttles were performed at the same cadence. The cadence was approximately 95 steps·min^{-1} for slow walking, walking with bags and floor sweeping; approximately 115 steps·min^{-1} for brisk walking; 130-140 steps·min^{-1} for slow running; and 145-160 steps·min^{-1} for fast running. Low jumps (approximately 2-5 cm) and higher jumps (> 5 cm) were performed continuously (one jump per second) for 20 s. Finally participants dropped from a 20 cm high box (typical stair step height) onto the floor, paused for a couple of seconds, returned to the top of the box, paused for a couple of seconds and repeated until they had completed ten box drops. No restrictions were placed on arm movement throughout activities.

Activity Monitors: A waveform GENEActiv (Gravity Estimator of Normal Everyday Activity, ActivInsights, Cambridgeshire, UK; seismic acceleration sensor, dynamic range +/- 8g) accelerometer and an ActiGraph GT3X+ (ActiGraph LLC, Pensacola, FL, USA; monolithic

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

differential capacitance sensor, dynamic range +/- 6g) accelerometer were worn over the right hip on an elastic waist belt. Both devices were tightly held together using tape to ensure they were subjected to the same accelerations. GENEActiv software (version 2.1) and ActiLife5 LITE analysis software were used to initialise the GENEActiv and the ActiGraph GT3X+, respectively, at a sampling frequency of 100 Hz and to upload the data. The two activity monitors were time-synchronized, i.e. initialized using the same computer clock. After completion of all activities, accelerometer data were uploaded on a computer and exported into Matlab (R2012b) for processing, using a custom graphical user interface.

Data processing:

From herein, bolded and non-bolded characters denote vectors and scalars respectively. The indexes GENE and GT3X+ refer to the GENEActiv and ActiGraph GT3X+ devices respectively. Both GENE and GT3X+ accelerometers output a [n_t x 3] array of values, where n_t = the value for each time point (t) and 3 = the number of acceleration axes (x, y and z), thus the array corresponds to the time histories of the three components of acceleration (a): $a_X(t)$, $a_Y(t)$ and $a_Z(t)$, t = 1 to n_T. From these the total acceleration vector **a(t)** was computed for each monitor as:

$$\mathbf{a}(\mathbf{t}) = a_X(t)\mathbf{x} + a_Y(t)\,\mathbf{y} + a_Z(t)\mathbf{z}\,, \text{(eq. 1)}$$

where **x**, **y**, **z** are the unit vectors of the accelerometers' local coordinate systems. The magnitude of acceleration a(t) was then computed:

$$a(t) = \sqrt{a(\mathbf{x})^2 + a(\mathbf{y})^2 + a(\mathbf{z})^2} \text{ (eq. 2)}$$

The time history of the magnitude was then graphed for both accelerometers. From these graphs, the magnitude peak corresponding to the first step of the first walking trial was manually identified for the two accelerometers. This allowed the synchronization of the time-stamped data

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

from the two devices to be confirmed. The datapoints corresponding to the beginning and end of each activity block were then manually identified on the graph for one monitor axis and the corresponding datapoints selected automatically for all axes of both monitors. For each block, one shuttle was randomly selected, except for jumping activities where the entire block was analysed. Finally, within each shuttle, the main peaks of acceleration associated with each step or jump were manually identified. The corresponding peak magnitudes are denoted ${}^{PEAK}{}_{i}a$, where i = 1,…,n denotes the i^{th} peak. The first and last peaks defined the start frame t_{START} and end frame t_{END} of each activity. Manual identification of each step/jump enabled automatic peak detection for steps/jumps.

The following features were calculated for each activity for each acceleration component $a_X(t)$, $a_Y(t)$, $a_Z(t)$ as well as for the total acceleration magnitude $a(t)$:

1. Time domain features from the entire signal:

a) Average magnitude:

$$^{AVG}a = \frac{\sum_{t=t_{START}}^{t_{END}} a(t)}{t_{END} - t_{START}},$$

b) Average of the absolute value with offset (gravity) removed:

$$^{AVG_ABS}a = \frac{\sum_{t=t_{START}}^{t_{END}} |a(t) - {}^{AVG}a|}{t_{END} - t_{START}},$$

c) the 25, 50 and 75th percentile of the absolute value with offset removed: ${}^{25}a$, ${}^{50}a$, ${}^{75}a$, respectively;

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

d) Root mean square (RMS) of the absolute value with offset removed:

$$\mathrm{RMS_ABS}_a = \sqrt{\frac{\sum_{t=t_{START}}^{t_{END}} [a(t) - {}^{AVG}a]^2}{t_{END} - t_{START}}},$$

2. Time domain features from the peaks in the signal:

a) Average peak magnitude:

$$\mathrm{AVG_PEAK}_a = \frac{\sum_{i=1}^{n} {}^{PEAK}{}_i a}{n},$$

b) Standard deviation (SD) of peak magnitude:

$$\mathrm{SD_PEAK}_a = \sqrt{\frac{\sum_{i=1}^{n} \left[{}^{PEAK}{}_i a - {}^{AVG_PEAK}a\right]^2}{n}},$$

c) The ratio between the average peak magnitude and the average of absolute value with offset removed:

$$\mathrm{PR} = \frac{\mathrm{AVG_PEAK}_a}{\mathrm{AVG_ABS}_a},$$

3. Frequency domain features from the signal:

a) An FFT was performed on the offset-removed acceleration, and the dominant frequency f_0 and its associated power P_0 were calculated. The FFT gives the frequency spectrum of the signal enabling the identification of underlying frequencies, or repeating patterns.

All the above variables were calculated for each acceleration component $a_X(t)$, $a_Y(t)$, $a_Z(t)$ as well as for the total acceleration magnitude $a(t)$. These features were selected as they have been used in previous pattern recognition studies (e.g. 3, 9, 11, 16, 23).

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Data analysis:

Descriptive statistics were calculated for all variables. A series of fully repeated measures ANOVAs (monitor x activity) were run to examine differences in each of the features across brand of monitor and across activities. The Greenhouse-Geisser correction was used where the assumption of sphericity was not met. Post-hoc analyses were carried out using pairwise comparisons with alpha (0.05) adjusted using the Bonferroni correction. To examine the equivalence of the dominant frequency f_0 identified for each activity, the proportion of participants with perfect agreement for the dominant frequency was calculated for each activity. Analyses were run for each acceleration component and total acceleration magnitude separately.

Part 2: Free-living

Part 2 investigated the equivalence of output from the GT3X+ and the GENE worn at the hip in children in a free-living setting. As the devices were compared in a "real life" setting the data for part 2 were high-pass filtered before analysis as has been carried out previously for free-living data (e.g. 18). The data were taken from a larger study investigating the relationship between psychological well-being, physical activity and physical activity context in children.

Participants: Fifty-eight boys and girls, aged 10-12 y (age: 10.7 ± 0.8 y, mass: 43.7 ± 10.8 kg, height: 1.49 ± 0.07 m), were recruited from schools in South Australia. The institutional ethics committee granted approval. Written informed consent and assent were obtained from the parents/guardians and children, respectively. Height was measured to the nearest 0.1 cm and body mass to the nearest 0.1 kg.

Data collection: Each participant wore two accelerometers at the hip (GT3X+, GENE) for seven consecutive days. The two hip monitors were taped securely together and positioned

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

above the right hip, on a belt worn round the waist. Children were instructed to wear the hip monitor night and day, and to only remove the hip monitor for water-based activities. The GT3X+ was initialised using Actilife version 6.5.3 and programmed to collect data at 80 Hz. The GENEs were initialised using GENEActiv PC software version 2.2 and programmed to collect data at 85.7 Hz. These were the maximum sampling frequencies that enabled data collection for >7 days. The small difference in sampling frequency between monitors will not have impacted on the output analysed due to the nature of the signal processing. All monitors were programmed to start collecting data at midnight on the day the child was given the monitors and time synchronised, i.e. initialized using the same computer clock.

Activity monitors: GT3X+ data were uploaded using Actilife version 6.5.3 and the .gt3x files were converted to .csv files containing measured acceleration data. The date and time stamp for each epoch was calculated from the file start time and sampling frequency. GENE data were uploaded using GENEActiv PC software version 2.2 and the .bin files were converted to .csv files containing the time and date stamp for each epoch and the measured acceleration data for each of the three axes. The total acceleration magnitude a(t) was calculated (see eqs. (1) and (2) above) and the offset (mainly due to gravity) was removed by applying a first-order high-pass filter (0.3 Hz cut-off frequency). The cut-off frequency was selected on the presumption that most acceleration related to human movement would occur at frequencies greater than 0.3 Hz; van Hees et al. (18) reported that the optimal cut-off frequency is probably between 0.2 and 0.5 Hz.

Data processing: Data from each monitor were run through a custom MATLAB code to elicit the following output for each participant for each monitor (GENE hip, GT3X+ hip):

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

a) Average of the high pass filtered total acceleration magnitude per day (g); $^{AVG_HPF}a$

Daily minutes (number of datapoints in acceleration range X sampling rate (Hz) X 60) accumulated at accelerations between 0-0.5 g; 0.5-1.0 g; 1.0-1.5 g; 1.5-2.0 g; >2.0 g;

b) $^{MIN}a_{0–0.5}$

c) $^{MIN}a_{1·0–1.5}$

e) $^{MIN}a_{1·5–2.0}$

f) $^{MIN}a_{>2.0}$

g) a plot of the time history of each acceleration component for each day.

Examination of the scatterplots between the $^{AVG_HPF}a$ measured by the GENE and the GT3X+ at the hip revealed three clear outliers (values greater than 2 SD away from the mean and/or residuals greater than 2 SD from the mean). Further examination of the data suggested technical issues with these GENE monitors and the three participants were excluded from all analyses reducing the sample size to 55.

The contemporaneous measured acceleration plots for each monitor for each participant for each day were carefully visually examined to determine which monitors could be matched for wear-time for each participant. Of the 55 participants, one participant did not wear any of the monitors during the measurement period, and two had large periods of sporadic non-wear during day-time (awake time), all three were excluded from analysis.

The outputs from the two hip monitors were matched for wear-time (day time only) for seven days in 50 participants, four days in one participant and two days in one participant resulting in a sample size of 50 for day-by-day analyses and 52 for analyses focusing on the

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

mean of measured days. The two participants with fewer than seven complete days had several days of non-wear or reduced wear (significant periods of non-wear during the waking day).

Data analysis:

Descriptive statistics were calculated for all output variables. A series of fully repeated measures ANOVAs (device × day) were carried out to test for differences in $^{AVG_HPF}a$ and time recorded in each of the acceleration bands between the GENE and the GT3X+ worn at the hip. Pearson's correlations were used to examine the relationships between the monitors for each of the output variables averaged across the seven days A fully repeated measures ANOVA (device × acceleration band) was used to test for differences in the pattern of accelerations measured by each accelerometer, averaged across all measured days. The Greenhouse-Geisser correction was used where the assumption of sphericity was not met. Post-hoc analyses were carried out using pairwise comparisons with alpha (0.05) adjusted using the Bonferroni correction.

Alpha was set at 0.01 due to the large number of statistical tests undertaken. IBM SPSS version 20.0 (IBM, Armonk, NY) was used for all statistical analyses.

Results

Part 1

Total acceleration magnitude

There was a significant main effect of activity for all features extracted from the time domain of the signal (e.g. Figures 1a-c), the peaks' amplitudes (Figures 2a-b) and from the frequency domain of the signal (Figures 2c-d). In general the magnitude of the features extracted

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

increased with ambulatory speed and with jump height. Values for box drops were relatively low for features based on the whole acceleration signal rather than peaks, reflecting the intermittent nature of the box drops (Figures 1a-c). However, peak amplitudes $^{AVG_PEAK}a$ and peak-to-average ratio was highest for box drops (Figures 2a and b).

The magnitude of the output from the GENE was generally higher than from the GT3X+ as evidenced by a significant main effect and interaction for all features except, ^{25}a, ^{50}a, ^{75}a (data not shown) and the dominant frequency f_0 (Figure 2c). The difference between monitors was significant for all activities for $^{AVG_ABS}a$ (Figure 1b), $^{RMS_ABS}a$ (Figure 1c), $^{AVG_PEAK}a$ (Figure 2a), $^{SD_PEAK}a$ (data not shown) and PR (Figure 2b). The ^{AVG}a from the GENE was significantly higher than from the GT3X+ for all activities other than the three walking activities and box drops (Figure 1a). The power of the dominant frequency P_0 from the GENE was significantly higher than from the GT3X+ for all activities other than walking and low jumps (Figure 2d).

The significant device x activity interactions were largely due to the difference in magnitude between the monitors' output increasing as the acceleration increased due to changes in speed and/or jump height (Figures 1a-c, 2c-d). However, this was not evident for the ^{25}a, ^{50}a, or ^{75}a where no main effect of monitor or device x activity interaction was present (data not shown). The differentiation between activities from the various features extracted was similar for the two monitors, despite the significant interactions evident for most features, with the pattern of significant results identical for the two monitors for ^{AVG}a (Figure 1a), $^{AVG_ABS}a$ (Figure 1b), the ^{25}a, ^{50}a, ^{75}a, f_0 (Figure 2c), p_0, (Figure 2d), and very similar for the remaining features ($^{RMS_ABS}a$ (Figure 1c), $^{AVG_PEAK}a$ (Figure 2a), $^{SD_PEAK}a$, PR (Figure 2d). Dominant frequency

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

power P_0 is the feature that resulted in the best discrimination between activities, with all activities significantly different from all others, with the exception of walking, walking with bags and sweeping (Figure 2d).

Individual axes

The patterns of results for axes X, Y and Z were similar to the pattern for the total acceleration signal (Figures 1 and 2). Figures for all features for each axis are available as Supplemental Digital Content. See:

- Figure, SDC1, Time domain features from the entire acceleration signal for axis 1, mainly vertical, http://links.lww.com/MSS/A404
- Figure, SDC2, Time and frequency domain features from the peaks of the acceleration signal for axis 1, mainly vertical, http://links.lww.com/MSS/A405
- Figure, SDC3, Time domain features from the entire acceleration signal for axis 2, mainly medio-lateral, http://links.lww.com/MSS/A406
- Figure, SDC4, Time and frequency domain features from the peaks of the acceleration signal for axis 2, mainly medio-lateral, http://links.lww.com/MSS/A407
- Figure, SDC5, Time domain features from the entire acceleration signal for axis 3, mainly antero-posterior, http://links.lww.com/MSS/A408
- Figure, SDC6, Time and frequency domain features from the peaks of the acceleration signal for axis 3, mainly antero-posterior, http://links.lww.com/MSS/A409

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Dominant frequency (f_0)

The number of participants with perfect agreement on the dominant frequency f_0 was 97-100% for over half of the activities (37/38 for walking, 38/38 for walking with bags, 37/38 for walking fast, 37/38 jogging fast and 37/38 for low jumps); and 87% or greater for the remaining activities (33/38 for sweeping, 34/38 for jogging slowly, 35/38 for high jumps) other than box drops, where agreement dropped to 75% (27/37). Agreement tended to be lower for individual axes, but followed a similar pattern.

Part 2

The $^{AVG_HPF}a$ values from the GENE were consistently higher than the values from the GT3X+ ($p<0.001$). However, the pattern of output across days was similar with no interaction evident (Figure 3) and the mean daily $^{AVG_HPF}a$ from the two monitors was positively correlated ($r = 0.93$, $p < 0.001$). The mean difference in $^{AVG_HPF}a$ between monitors was 0.010 g.

The number of minutes recorded in the lowest acceleration band (0-0.5 g) was higher for the GT3X+ than the GENE ($p<0.001$, Figure 4a), but higher for the GENE than the GT3X+ for all other acceleration bands (0.5-1.0 g, 1.0-1.5 g, 1.5-2.0 g, 2.0+ g, $p<0.01$, Figures 4b-e). However, the pattern across days was similar in all acceleration bands, with no significant interactions observed ($p>0.05$, Figures 4a-e).

The pattern of time recorded across acceleration bands was similar as shown by strong positive correlations between the mean daily time recorded by the two monitors in each acceleration band: 0-0.5 g, $r = 0.93$ $p < 0.001$; 0.5-1.0 g $r = 0.95$, $p<0.001$; 1.0-1.5 g, $r = 0.96$, $p < 0.001$; 1.5-2.0 g, $r = 0.97$, $p < 0.001$; 2.0+ g, $r = 0.68$, $p < 0.001$. There were small differences in

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

the number of minutes recorded resulting in a significant device x acceleration band interaction ($p<0.001$, Table 1).

Discussion

A key recommendation from the 2009 *Objective Measurement of Physical Activity Meeting* co-sponsored by the *National Institute of Health* and *American College of Sports Medicine* was that monitor data should be collected and saved as raw signals, with data transformation carried out post processing to facilitate comparisons between output regardless of which monitor is used (2, 4, 5, 15). This study compared the measured acceleration output under both laboratory and field conditions as a step towards determining whether and under which conditions the measured output from the GENE and the GT3X+ are interchangeable, as recommended by Welk et al. (14). Part 1 of this study examined multiple time and frequency domains from contemporaneous output from the GENE and the GT3X+ across a variety of specified activities. Part 2 of this study assessed whether the magnitude and pattern of accelerations from the GENE and the GT3X+ during free-living agreed and/or were related.

Part 1: Laboratory testing

As hypothesized, the amplitudes of accelerations were larger for the GENE than the GT3X+. This was true for both the total magnitude and the amplitude of each individual component. This supports and extends earlier preliminary findings in small samples that reported higher mean peaks of acceleration from the GENE than the GT3X+ during specific activities (10) and during mechanical shaker testing (6). For this reason, classification algorithms or intensity cut-points based on time domain features may not be directly transferrable from one

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

device to the other. Consistent with these findings, John et al. (6) reported that the prediction accuracy for classification of activity type in eight participants dropped ~8% ($p<0.05$) when applying a model based on time domain features from the GENE to the GT3X+, although the drop in accuracy (~2%) was not significant when applying a model based on time domain features from the GT3X+ to the GENE.

It is not clear why there appears to be a consistent difference in the magnitude of the output between the two brands of monitor. Possible reasons include the greater dynamic sensing capacities of the GENE (+/- 8 g) relative to the GT3X+ (+/- 6g), or that the two brands use different microelectromechanical systems (MEMS) sensors; the GENE uses the Analog devices ADXL345 and the GT3X+ uses the Kionix® (Ithaca, NY, USA) KXSC7-3672 accelerometer. Furthermore, while the measured data are not filtered by the GENE, the GT3X+ does apply filtering. The details of this filtering process are proprietary (6).

To investigate whether a correction factor could be used to convert between the time domain features from one device to the other, we assessed the linear fit of the average measured acceleration magnitude, the average measured acceleration magnitude with the offset (gravity) removed, and the average peak acceleration magnitude. Figure 5a presents the average measured acceleration magnitude measured by the GENE (Y-axis) against that measured by the GT3X+ (X-axis). It appears from this figure that the relationship between acceleration magnitudes reported by both devices is linear. The coefficient of determination from Pearson's correlation of the linear fit to ${}^{AVG}a_{GENE} = f({}^{AVG}a_{GT3X+})$ is $r^2 = 0.965$. When removing gravity from the signals, the GENE reports amplitudes 17% larger than the GT3X+ and the linearity is excellent with an r^2 of 0.977 (Figure 5b). Finally the mean peak amplitudes (gravity removed) also show a very good

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

linear relationship with an r^2 of 0.947, although heteroscedasticity is evident due to increased variability at higher accelerations (Figure 5c). These findings support the conclusions of John et al. (6) who reported an increasing difference in the mean acceleration from the GENE and GT3X+ as the magnitude of acceleration increased using a mechanical shaker. The linearity of the relationship was not assessed in that study; however, the mean values John et al. (6) presented for the x-axis of the GENE and GT3X+ during mechanical testing are almost perfectly linearly related ($R^2 = 0.999$, $p < 0.001$).

In conclusion, it should be possible to apply time-domain-based classification algorithms developed for one device to the other by simply applying an affine conversion on the measured acceleration values. The difference in acceleration magnitudes seems to be the only significant difference between the devices. In the frequency domain, as hypothesized, the two monitors agreed extremely well on the dominant frequency as well as on the power of that frequency. Once again this is in accordance with John et al. (6) who reported that prediction accuracy was not significantly compromised (less than 3% difference) when using a classification algorithm based on frequency-domain features on both devices.

The present study did not aim to develop classification algorithms. However, the examination of the different time- and frequency-domain features presented here may aid decisions as to which features to include in such algorithms. As stated previously, methods based purely on frequency domain features do not distinguish very well between activities having the same typical frequency but different intensity levels. For example, in the present study, it is not possible to differentiate jumps from a fast walk using the dominant frequency only, because the frequency of execution of both movements was similar. The same observation could be made for box drops and jogging, although these activities are different in terms of intensity. If the power

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

of the main frequency is added to the criteria, the distinction between box drops and jogs becomes very obvious (Figure 2d). Adding time-domain features may help refine the classification even further; for instance, the peaks' amplitudes (Figure 2a) enables separating a normal walk from a sweep, which is not possible from frequency domain analysis alone. Peaks' amplitudes also result in a better distinction between box drops and walks than frequency-domain features. Additionally, it is worth noting that the box drops in this study were performed in a rather periodic manner, which may not happen in real-life situations. In that case, time-domain features may help improve the classification even further for these types of non-cyclic activities where frequency analysis may fail. Thus, optimal classification performance across a range of activities would likely be obtained from a model including features from both the time and frequency domains.

Part 2: Field testing

As anticipated based on the laboratory results, the average daily acceleration from the GENE was higher than the GT3X+. However, the day to day difference was fairly consistent with the GT3X+ measuring 12-13% lower than the GENE and the output was highly correlated. Applying a correction factor to the data removes the significant difference between the two monitors (data not shown). It should be noted that this correction factor was developed on the same data set as it was applied to and correction factors would likely differ depending on the range and level of accelerations encountered. This difference in acceleration magnitude across monitor brand suggests intensity cut-points would need to be brand-specific, or a conversion factor available for switching between the two brands.

The high concurrent validity, albeit with a mean bias, of the two brands of accelerometer apparent in the free-living study suggests that the activity x monitor interactions evident in the

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

laboratory part of the study did not translate into inconsistent monitor differences during free-living measures in this population. This may be due to the differences between monitors being fairly consistent for walking activities and increasing for running and jumping. Only a relatively small portion of the day is spent in these higher intensity activities during free-living, even in children (1). Thus, the majority of measured time would be in the low intensity range, where values are more comparable, e.g. the average acceleration during walking was ~0.2 g, with peaks of up to 1 g. The difference in time recorded by the two monitor brands at 0-0.5 g was 0.4% (GENE: 1402.4±13.4 min; GT3X+: 1408.6±11.9 min) and the difference in time recorded at 0.5-1 g was 14% (GENE: 26.4±8.6 min; GT3X+: 22.7±7.1 min).

Limitations

The evaluation of monitor comparability in both a laboratory and field setting is a strength of this study. This enabled the comparison of specific features from the time and frequency domain of the signal across a range of different activities, and an assessment of the impact of these differences on the average daily acceleration and the time recorded in different bands of acceleration. Further, the detailed examination of which features differentiated between which activities and which features differed between monitor brands will inform the development of classification algorithms. However, the lack of common household, occupational or sports tasks in the laboratory protocol is a potential limitation.

A high pass filter was used to try and remove the effects of gravity from the acceleration signal. This method relies on assumptions about the frequency content of the gravitational component of the signal (13). We acknowledge that this method may not effectively remove the effects of gravity from the signal. Recently, van Hees et al. (13) evaluated five methods of removing the gravity component from the acceleration signal and found no single metric

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

outperformed all others. As the GT3X+ and GENE were treated identically, the method of processing should not impact on the results of the monitor comparison.

Conclusions

In conclusion, features of the frequency-domain compare well between monitors. Hence, frequency-domain-based classification algorithms should be transferable between monitors. In the time domain, although the magnitude of features is greater for the GENE than the GT3X+, there are strong linear relationships between the features from the two monitors. Therefore, it should be possible to apply classification algorithms based on time-domain or both time- and frequency-domain features, developed for one device to the other by simply applying an affine conversion on the measured acceleration values.

The strong relationship between mean daily accelerations measured by the two brands suggests the day-to-day pattern of habitual activity and ranking of activity level within a group will be similar irrespective of the brand used. Absolute values will differ; however application of an appropriate conversion factor should make values interchangeable between the two brands.

Acknowledgements: Adult data (part 1) were part of a larger study funded by an Innovative Award from the National Osteoporosis Society, UK. The authors would like to thank Pippa Griew, David Burdon, Grace Neate, Claire Stanley and Caitlin Stanley for assistance with data collection. No external funding was received for this study. The results of the present study do not constitute endorsement by the authors or the American College of Sports Medicine of the products described. There are no conflicts of interest.

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

References

1. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. Prev Med. 2007; 44: 143–147.

2. Basset DR, Rowlands AV, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc. 2012; 44: S32-S38.

3. Bonomi AG, Goris AH, Yin B, Westerterp KR. Detection of type duration and intensity of physical activity using an accelerometer. Med Sci Sports Exerc. 2009; 41: 1770–7.

4. Butte NF, Ekelund U, Westerterp KR. Assessing Physical Activity Using Wearable Monitors: Measures of Physical Activity. Med Sci Sports Exerc. 2012; 44: S5–S12.

5. Chen KY, Janz KF, Zhu W, Brychta RJ. Redefining the Roles of Sensors in Objective Physical Activity Monitoring. Med Sci Sports Exerc. 2012; 44: S13–S23.

6. John D, Sasaki J, Staudenmayer J, Mavilia M, Freedson PS. Comparison of raw acceleration from the GENEA and ActiGraph™ GT3X+ activity monitors. Sensors (Basel). 2013; 13: 14754-63.

7. Kavanagh JJ, Menz HB. Accelerometry: a technique for quantifying movement patterns during walking. Gait Posture. 2008; 28:1-15.

8. Oshima Y, Kawaguchi K, Tanaka S, Ohkawara K, Hikihara Y, Ishikawa-Takata K, Tabata I. Classifying household and locomotive activities using a triaxial accelerometer. Gait Posture. 2010; 31: 370–4.

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

9. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. Med Sci Sports Exerc. 2006; 37: 1626–34.

10. Rowlands AV, Stiles VH. Accelerometer counts and raw acceleration output in relation to mechanical loading. J Biomech. 2012; 45: 448-454.

11. Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson PS. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J Appl Physiol. 2009; 107: 1300–7.

12. Stiles VH, Griew PJ, Rowlands AV. Use of accelerometry to classify activity beneficial to bone in premenopausal women. Med Sci Sports Exerc. 2013; 45: 2353-61.

13. van Hees VT, Gorzelniak L, Dean León EC, Eder M, Pias M, Taherian S, Ekelund U, Renström F, Franks PW, Horsch A, Brage S. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One 2013; 23(8):e61691.

14. Welch WA, Bassett DR, Thompson DL et al. Classification Accuracy of the Wrist-Worn GENEA Accelerometer. Med Sci Sports Exerc. 2013; 45: 2012-9.

15. Welk G J, McClain J, Ainsworth BE. Protocols for Evaluating Equivalency of accelerometry-Based activity Monitors. Med Sci Sports Exerc. 2012; 44: S39–S49.

16. Zhang S, Murray P, Zilmer R, Eston RG, Catt, M, Rowlands AV. Activity classification using the GENEA: Optimum sampling frequency & number of axes. Med Sci Sports Exerc. 2012; 44: 2228-2234.

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

17. Zhang S, Rowlands AV, Murray P, Hurst TL. Physical activity classification using the GENEA wrist-worn accelerometer. Med Sci Sports Exerc. 2012; 44: 742-748.

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Figure legends

Figure 1. Time domain features from the entire acceleration signal: a) Average magnitude (^{AVG}a); b) Average of the absolute value with offset (gravity) removed ($^{AVG_ABS}a$; c) Root mean square (RMS) of the absolute value with offset removed ($^{RMS_ABS}a$).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

Figure 2. Time and frequency domain features from the peaks of the acceleration signal: a) Average peak magnitude $^{AVG_PEAK}a$; b) The ratio between the average peak magnitude and the average of absolute value with offset removed (PR). N.B. This latter figure has two panels: The upper panel shows the full range of ratios and the lower panel expands the axis to show the ratios for the first eight activities more clearly; c) dominant frequency (f_0); d) the power associated with the dominant frequency (p_0).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

Figure 3. Average of the high pass filtered total acceleration magnitude per day ($^{AVG_HPF}a$) for the GENE (solid line) and GT3X+ (dashed line).

Figure 4. Daily minutes recorded by the GENE (solid line) and GT3X+ (dashed line) in acceleration bands : a) 0-0.5 g, b) 0.5-1.0 g, c) 1.0-1.5 g, d) 1.5-2.0 g, e) >2.0 g.

Figure 5. Linear fit of a) (^{AVG}a), b) $^{AVG_ABS}a$) and c) $^{AVG_PEAK}a$;) measured by the GENE (Y-axis) against that measured by the GT3X+ (X-axis). Dashed line = line of identity.

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

List of Supplementary Digital Content

SDC 1. Time domain features from the entire acceleration signal for axis 1, mainly vertical: a) Average magnitude (^{AVG}a); b) Average of the absolute value with offset (gravity) removed ($^{AVG_ABS}a$; c) Root mean square (RMS) of the absolute value with offset removed ($^{RMS_ABS}a$).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

SDC 2. Time and frequency domain features from the peaks of the acceleration signal for axis 1, mainly vertical: a) Average peak magnitude ($^{AVG_PEAK}a$; b) The ratio between the average peak magnitude and the average of absolute value with offset removed (PR). N.B. This latter figure has two panels: The upper panel shows the full range of ratios and the lower panel expands the axis to show the ratios for the first eight activities more clearly; c) dominant frequency (f_0); d) the power associated with the dominant frequency (p_0).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

SDC 3. Time domain features from the entire acceleration signal for axis 2, mainly medio-lateral: a) Average magnitude (^{AVG}a); b) Average of the absolute value with offset (gravity) removed ($^{AVG_ABS}a$); c) Root mean square (RMS) of the absolute value with offset removed ($^{RMS_ABS}a$).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

SDC 4. Time and frequency domain features from the peaks of the acceleration signal for axis 2, mainly medio-lateral: a) Average peak magnitude ($^{AVG_PEAK}a$; b) The ratio between the average peak magnitude and the average of absolute value with offset removed (PR). N.B. This latter figure has two panels: The upper panel shows the full range of ratios and the lower panel

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

expands the axis to show the ratios for the first eight activities more clearly; c) dominant frequency (f_0); d) the power associated with the dominant frequency (p_0).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

SDC 5. Time domain features from the entire acceleration signal for axis 3, mainly antero-posterior: a) Average magnitude (^{AVG}a); b) Average of the absolute value with offset (gravity) removed ($^{AVG_ABS}a$); c) Root mean square (RMS) of the absolute value with offset removed ($^{RMS_ABS}a$).

* significant difference between monitors within activity ($p < 0.006$)

† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

SDC 6. Time and frequency domain features from the peaks of the acceleration signal for axis 3, mainly antero-posterior: a) Average peak magnitude ($^{AVG_PEAK}a$; b) The ratio between the average peak magnitude and the average of absolute value with offset removed (PR). N.B. This latter figure has two panels: The upper panel shows the full range of ratios and the lower panel expands the axis to show the ratios for the first eight activities more clearly; c) dominant frequency (f_0); d) the power associated with the dominant frequency (p_0).

* significant difference between monitors within activity ($p < 0.006$)

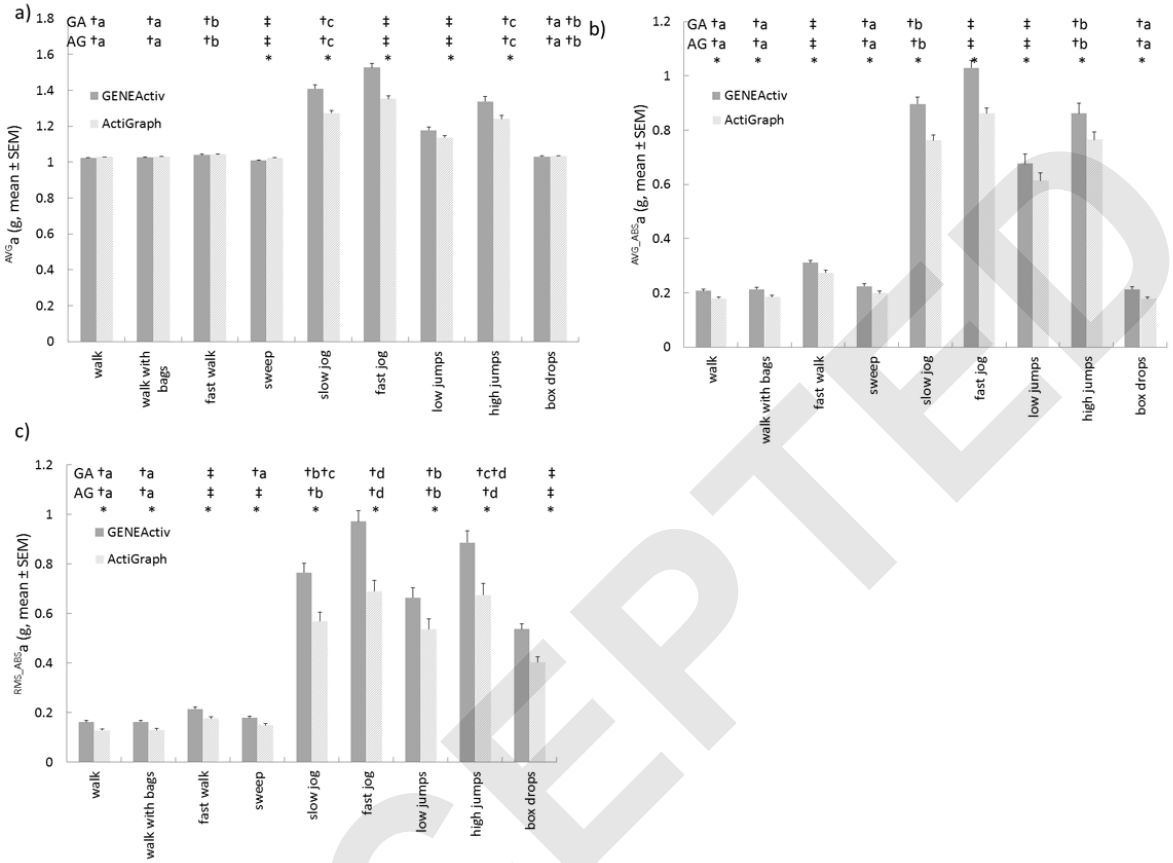
† not different from activities with the same letter, but different from all other activities (within monitor where GA or AG stated (GA = GENE, AG = GT3X+))

‡ different from all other activities (within monitor where GA or AG stated)

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

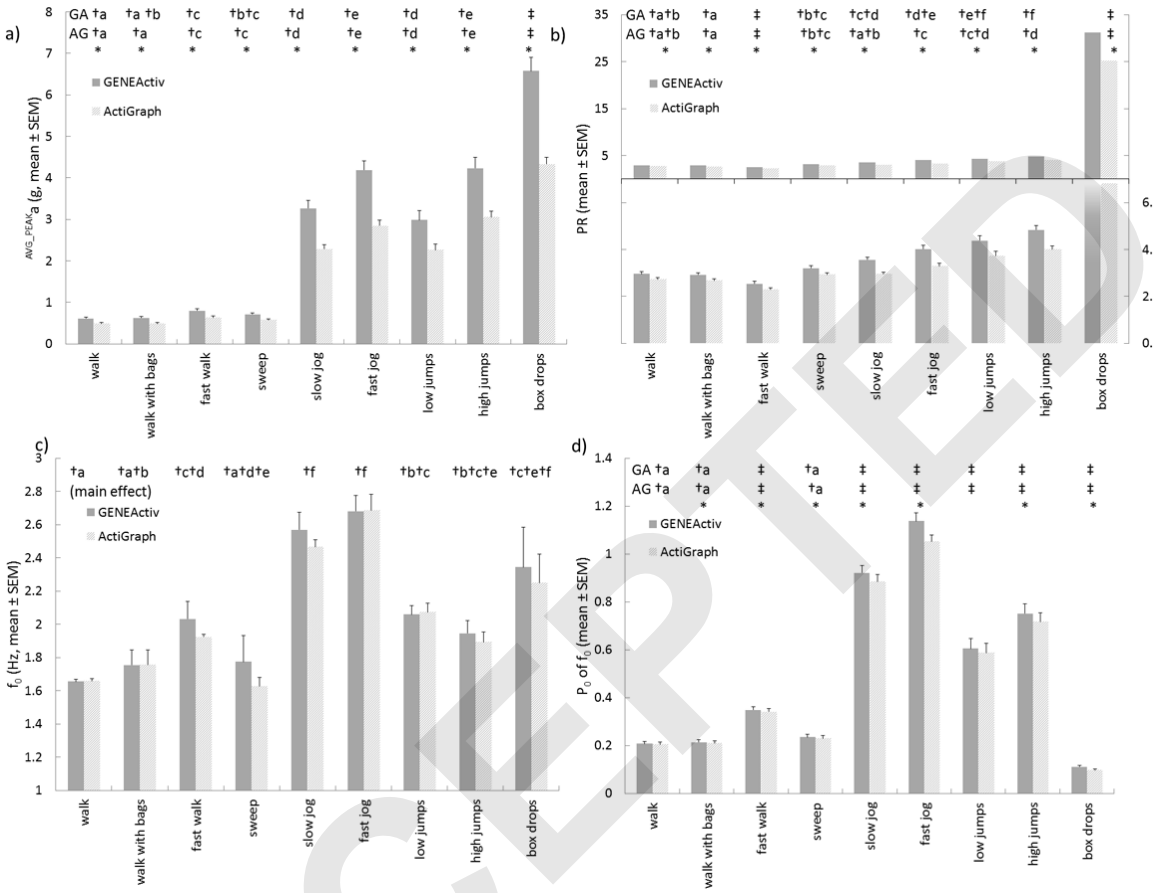
Figure 1



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

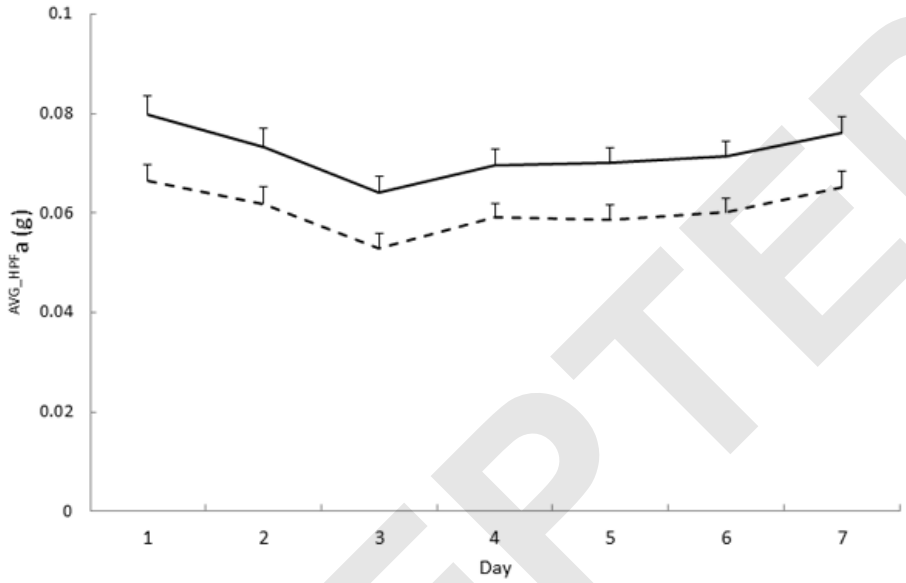
Figure 2



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

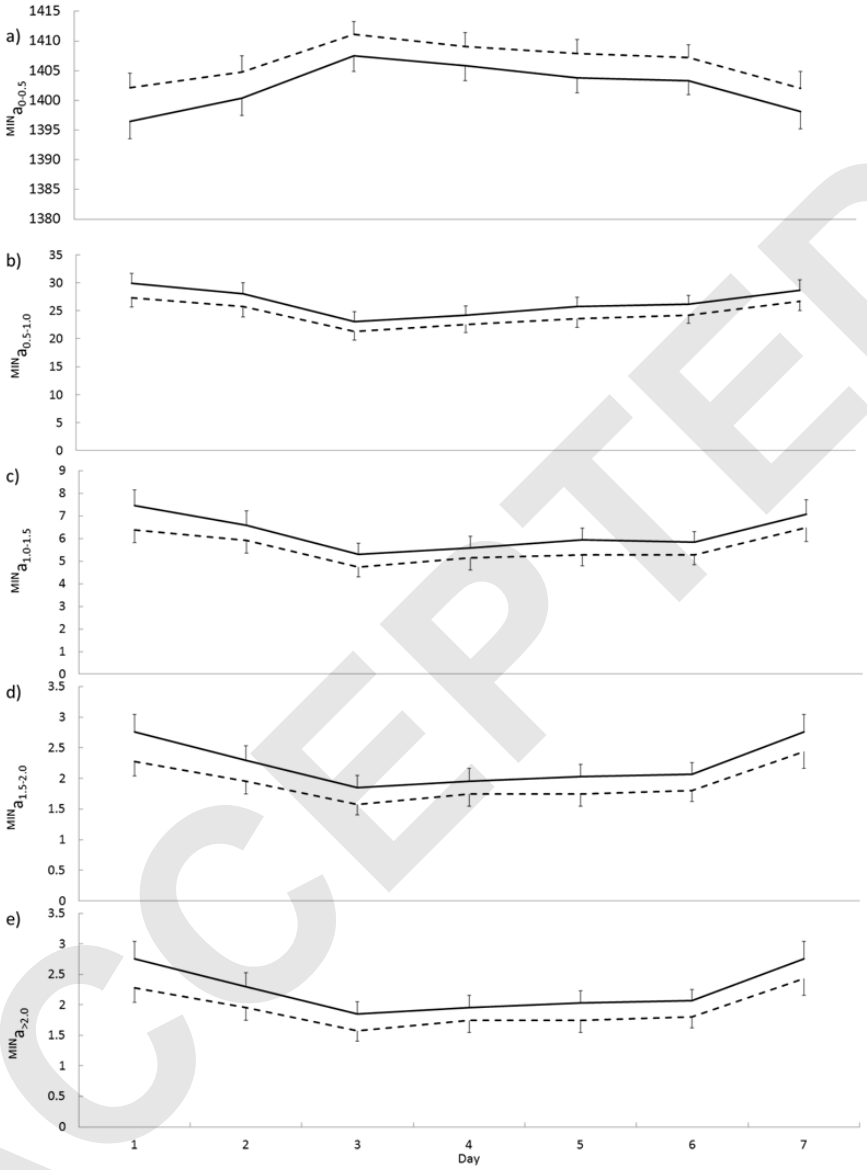
Figure 3



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

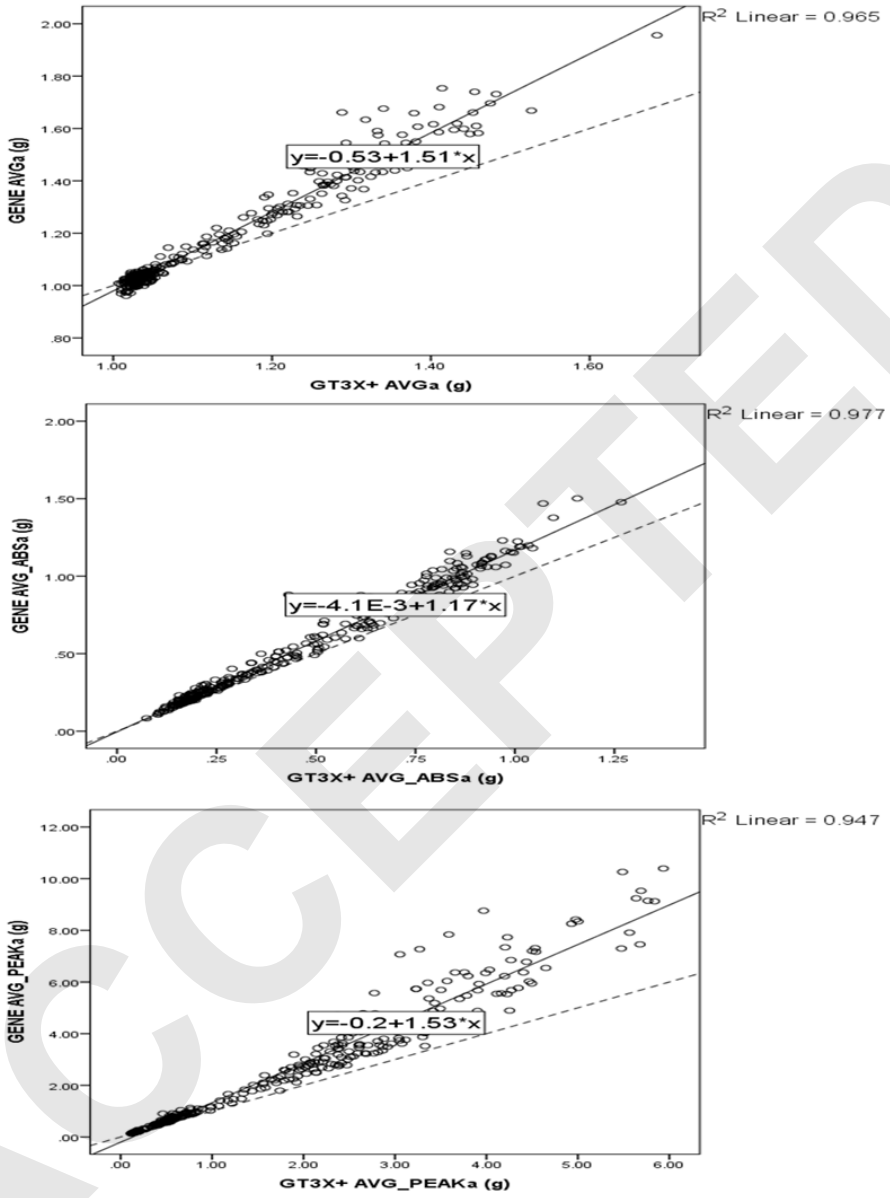
Figure 4



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Figure 5



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

Table 1. Amount of time recorded in each acceleration band by monitor brand (GENE and GT3X+).

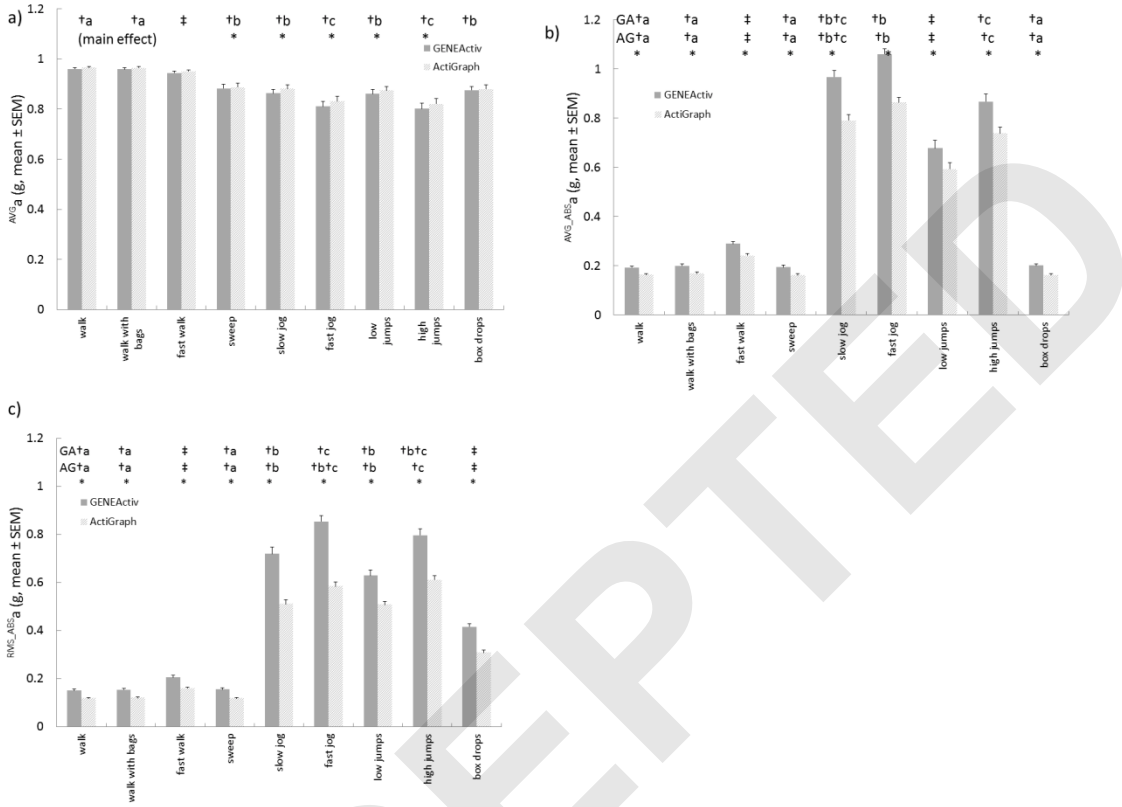
	Mean (SD) daily time recorded in acceleration band (min)	
Acceleration band (g)	GENE	GT3X+
0-0.5	1402.4 (13.4)†	1408.6 (11.9)
0.5-1.0	26.4 (8.6)†	22.7 (7.1)
1.0-1.5	6.2 (2.7)†	5.3 (2.5)
1.5-2.0	2.2 (1.1)†	1.8 (1.0)
>2.0	2.6 (1.8)†	1.5 (0.9)

†significant difference between monitor/wear-site ($p<0.001$).

ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

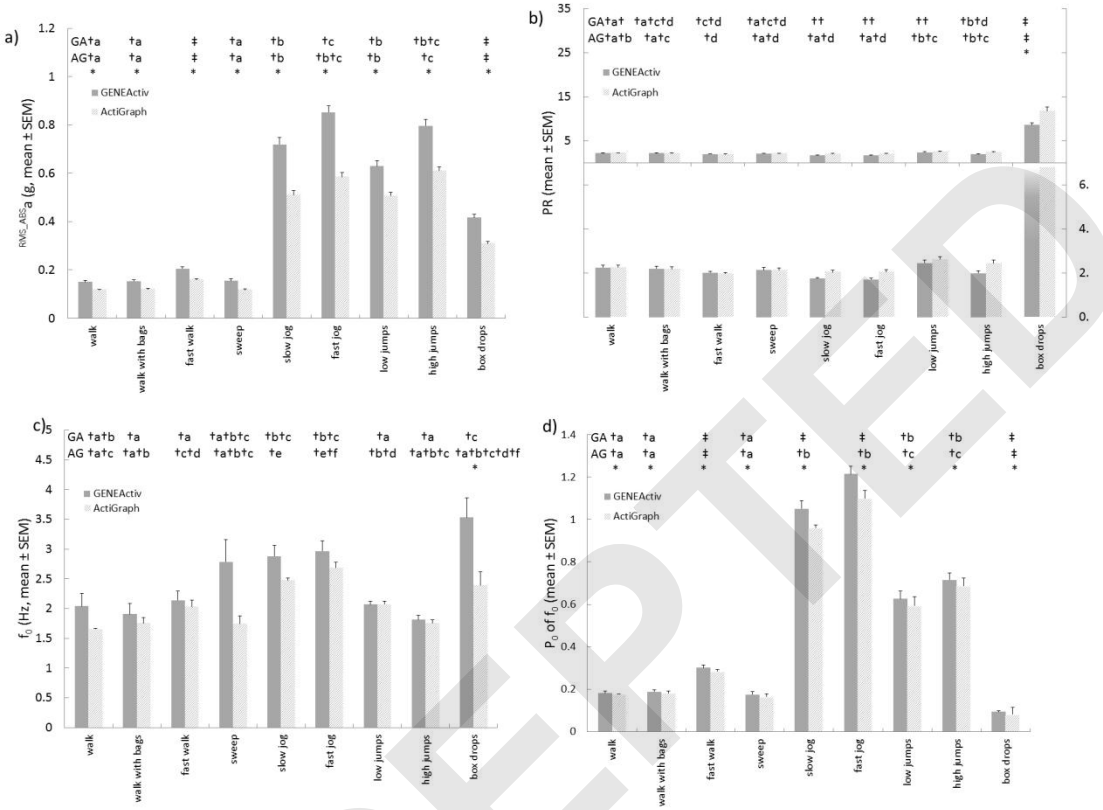
SDC F1



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

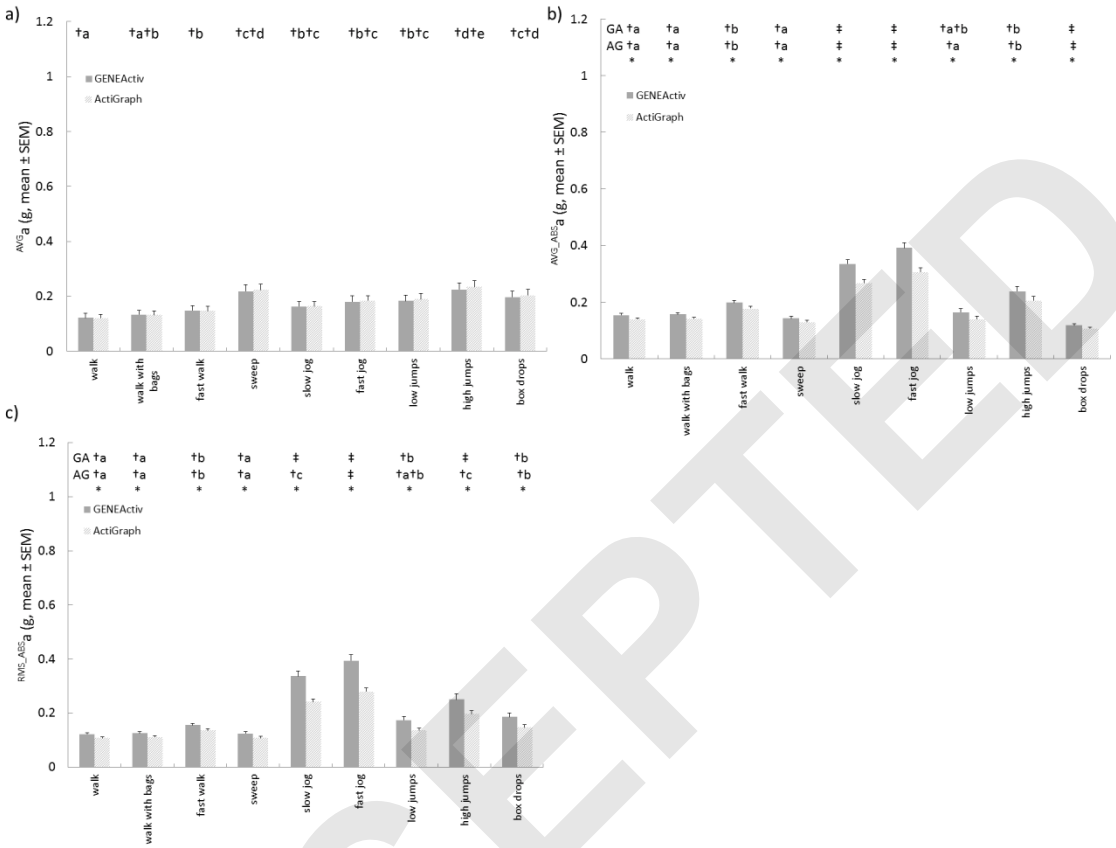
SDC F2



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

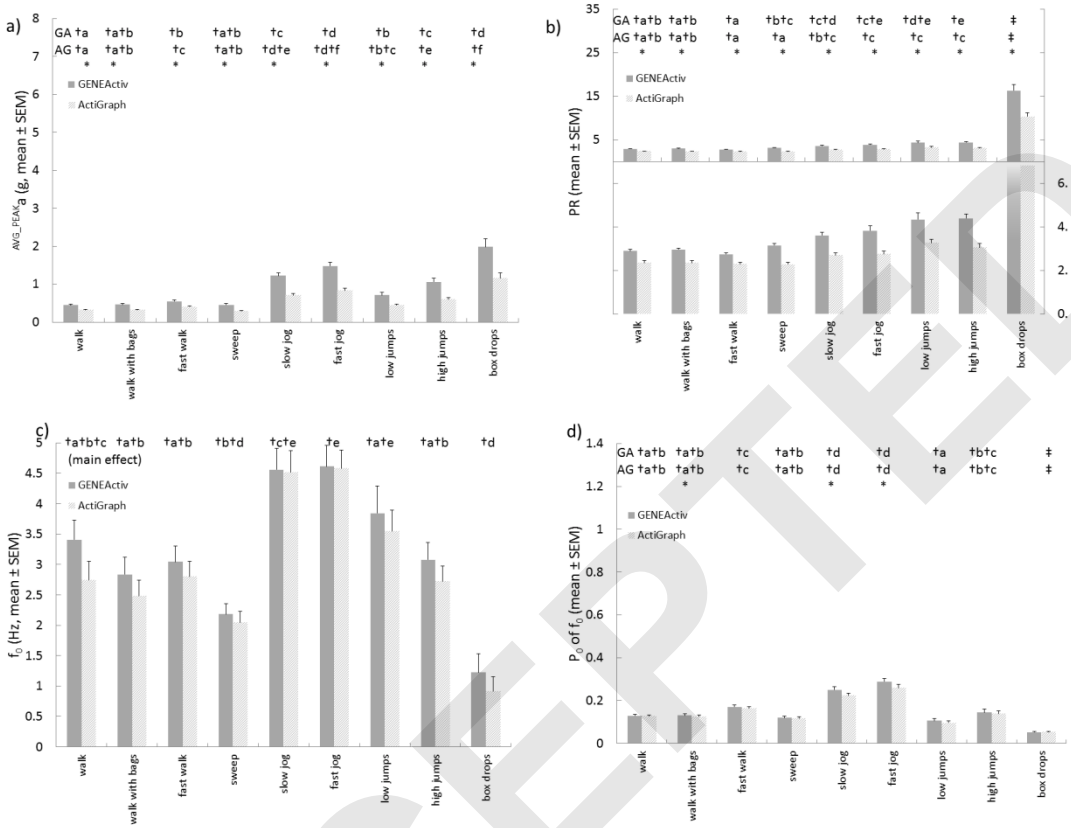
SDC F3



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

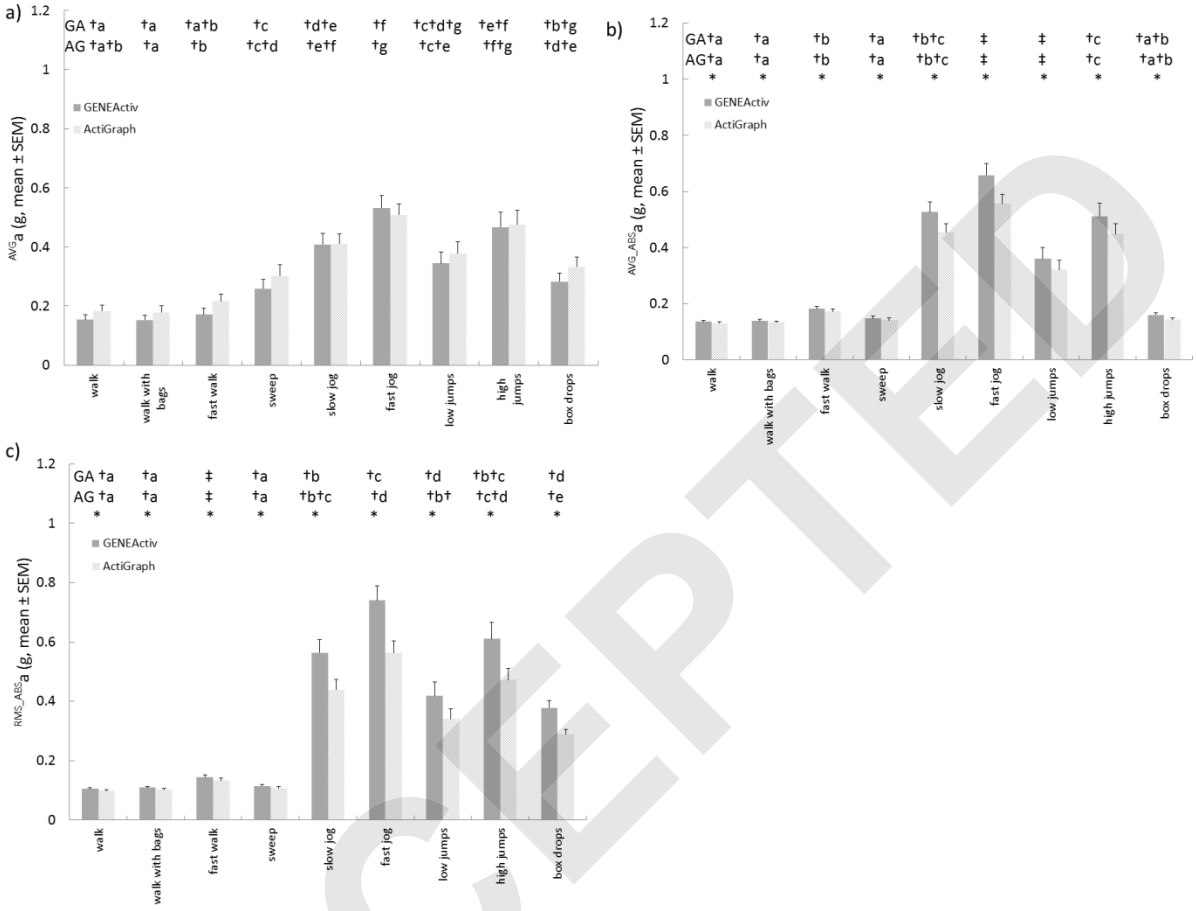
SDC F4



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

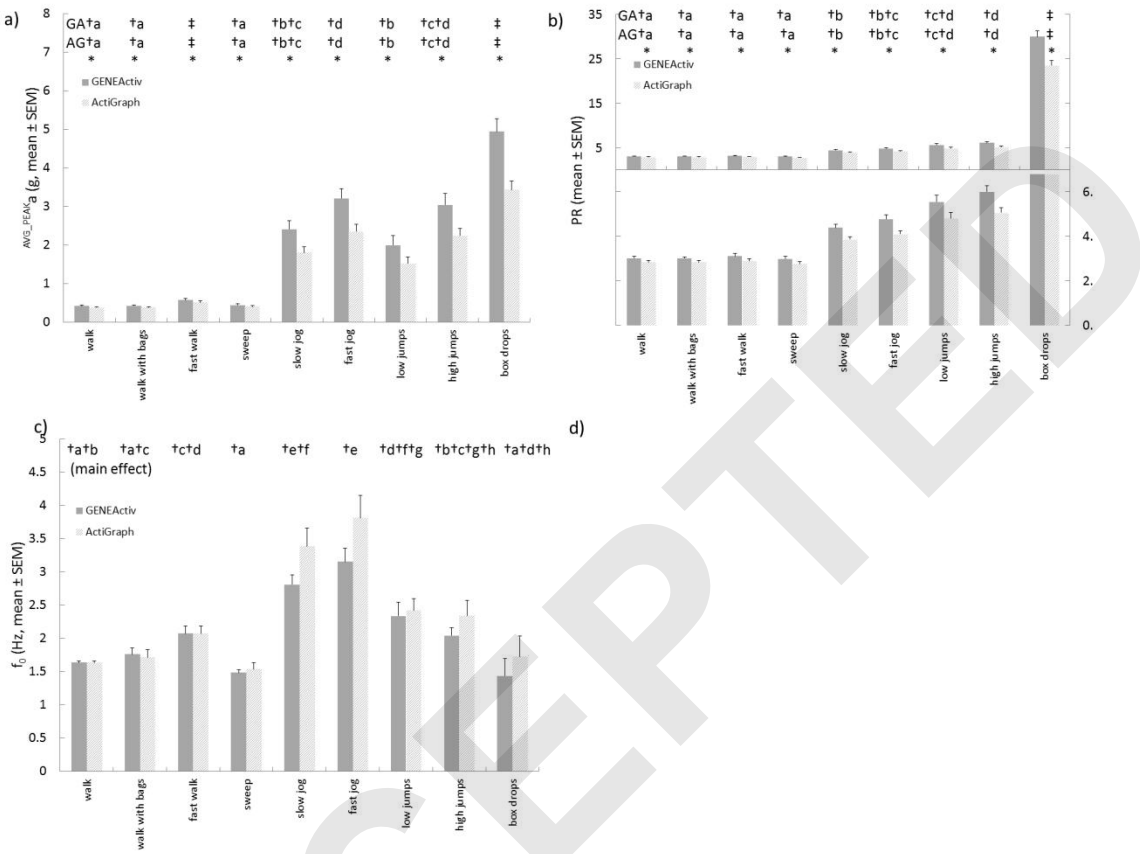
SDC F5



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.

SDC F6



ACCEPTED

Copyright © 2014 by the American College of Sports Medicine. Unauthorized reproduction of this article is prohibited.