Interpretation of associations between the accelerometry physical activity spectrum and cardiometabolic health and locomotor skills in two cohorts of children using raw, normalized, log-transformed, or compositional data

Eivind Aadland (eivind.aadland@hvl.no),[1] Lars Bo Andersen (lars.bo.andersen@hvl.no),[1] Jairo Hidalgo Migueles (jairohm@ugr.es),[2] Francisco B Ortega (ortegaf@ugr.es),[2,3] Olav Martin Kvalheim (olav.kvalheim@uib.no)[4]

[1]Western Norway University of Applied Sciences, Faculty of Education, Arts and Sports, Department of Sport, Food and Natural Sciences, Campus Sogndal, Sogndal, Norway

[2]PROFITH "PROmoting FITness and Health through physical activity" Research Group, Sport and Health University Research Institute (iMUDS), Department of Physical and Sports Education, Faculty of Sport Sciences, University of Granada, Granada, Spain

[3]Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden.

[4]University of Bergen, Department of Chemistry, Bergen, Norway

Corresponding author

Eivind Aadland

Western Norway University of Applied Sciences, Faculty of Education, Arts and Sports, Department of Sport, Food and Natural Sciences, Campus Sogndal, Box 133, 6851 Sogndal, Norway.

Phone: +47 5767 6086; Email: eivind.aadland@hvl.no

Word count main text: 5044; abstract: 199

Abstract

It is discussed whether associations between accelerometer-derived physical activity intensities and outcomes should be analyzed as absolute or relative data. The aim of the present study was to compare interpretation of association patterns of spectrum physical activity descriptions with outcome using raw, normalized, log-transformed, or compositional data. We used two datasets including 1) 841 schoolchildren and a cardiometabolic health outcome and 2) 1081 preschool children and a locomotor skill outcome. Accelerometry (ActiGraph GT3X+) data were described using multiple variables across the intensity spectrum. We varied the binning of variables to examine sensitivity of the compositional analyses to changes in the distribution center. We used multivariate pattern analysis for all analyses and interpretations of data. Analyses of absolute (i.e., non-compositional) data showed weak associations for lower intensities and strongest associations with cardiometabolic health and locomotor skills for vigorous intensities. The same association patterns were partly observed for the compositional data, but association patterns were in some cases conflicting. The binning of variables had a major influence on associations for compositional data, but not for absolute data, meaning that conclusions depend on the operationalization of compositional data. These differences challenge and confuse interpretation of association patterns derived from the different approaches.

Keywords Multivariate pattern analysis; Compositional data analysis; Multicollinearity; Singularity, Log-transform, Log-ratio; Children; Accelerometer

Background

Accelerometers capture movement on an absolute scale, from which time spent in different intensities is commonly derived. If defining a measurement period as finite, for example the wear period or the 24 hours of a day, movement behaviors can be considered compositional. According to this view, each behavior only contains information relative to the whole composition [1, 2]. Thus, within a fixed period, varying time spent in different activities will substitute each other and be reallocated, for example along the physical activity (PA) intensity spectrum. This feature of "closure" has recently stimulated discussion about the best way to handle and analyze accelerometer-derived PA data described by multiple correlated variables [3]. Both isotemporal substitution models [4] and compositional analysis [1, 2] have been suggested as possible approaches to handle this closure or constant-sum constraint (see Aadland et al [3] for a brief review of the methods). However, both approaches commonly rely on linear least squares regression analysis as their statistical underpinning. Thus, although these approaches theoretically treat PA data as compositional, and thus solve the closure problem with respect to reallocation of time across variables, linear regression cannot handle the much greater challenge of strong multicollinearity resulting from activity patterns inducing correlation patterns among PA variables. This multicollinearity makes linear regression analysis inappropriate for analysis of such data [5, 3], already with few explanatory variables.

The deficiency of linear regression is further exaggerated by the use of more detailed PA descriptions as compared to traditional blunt descriptions typically using 3–4 variables [3], for example sedentary time (SED), light PA (LPA), moderate-to-vigorous PA (MVPA), and/or sleep. In line with recent suggestions for an improved use of accelerometer data [6-8], Aadland et al recently showed that the inclusion of the entire intensity spectrum and using an improved resolution of the PA description greatly increased information derived from accelerometry, and thus knowledge about associations between PA and cardiometabolic health, in children [9, 10]. These findings show the unexploited potential of accelerometry when moving from summary measures to more detailed descriptions of

PA. Yet, such data requires statistical methods that are able to handle this type of highly correlated data. Therefore, Aadland et al introduced multivariate pattern analysis to analyze associations between the multicollinear explanatory PA variables and the cardiometabolic health outcomes [9, 10]. Multivariate pattern analysis is widely applied in other fields of research with the objective of revealing patterns of important biomarkers among hundreds or thousands of highly interrelated variables [11-13], and can handle completely collinear explanatory variables using latent variable modelling [14, 15]. Since multivariate pattern analysis can handle multicollinear data, Aadland et al [3] were able to circumvent the limitations of linear regression and directly compare association patterns between raw (min/day) and compositional PA data with cardiometabolic health. Both a traditional description reducing the PA profiles to four variables (SED, LPA, moderate PA (MPA), and vigorous PA (VPA)) and a spectrum description of 23 variables were used. With both descriptions we found apparently differing association patterns with cardiometabolic health using raw or compositional data. We also found that models using compositional data lead to higher explained variances than models using raw data, which could be attributed to reduction of noise due to log-transformation [16]. However, it was not examined whether the improved model fit resulted from the log-transformation or from the centering of data. Thus, it should be examined whether the improved model fit is a favorable feature of compositional data compared to raw and/or log-transformed data.

We did not provide a comprehensive elaboration on interpretation of the findings from raw or compositional data in the previous paper [3], which is crucial to fully understand the results and thus inform PA guidelines and provide direction for future research. Thus, a broader exploration and discussion of the interpretation of the absolute (raw, normalized, and log-transformed data) and relative (compositional data) nature of these association patterns are needed. The aim of this paper is therefore to extend our previous analysis [3] on comparison of results and interpretation of associations using raw, normalized, log-transformed, and compositional PA data. For improved generalization, we will analyze two large datasets (in preschool- and schoolchildren) and two different outcomes (cardiometabolic health and locomotor skills) in this work.

Methods

We have previously published the PA signature associated with cardiometabolic health in the *Active Smarter Kids (ASK)* study [9, 10, 3] and the PA signature associated with locomotor skills in *The Sogn og Fjordane Preschool Physical Activity Study (PRESPAS)* [17]. The aim of the present study is limited to *compare* association patterns using raw and compositional data within these datasets. We refer readers to previously published descriptions of sampling and children's characteristics, study protocols, instruments, and procedures of the ASK study [9, 10, 3, 18] and the PRESPAS study [19, 17] for detailed study information. Thus, we provide below only a brief overview of the most relevant information to provide sufficient context to support the study aim of comparing association patterns between models relying on different treatment of the explanatory data matrix.

Participants

The ASK study was conducted in western Norway during 2014–2015 and included 841 10-year old schoolchildren providing relevant explanatory (PA) and outcome (cardiometabolic health) data [9, 10, 3, 18]. The PRESPAS study was conducted in western Norway during 2015–2016 and included 1081 3-6-year old preschool children providing relevant explanatory (PA) and outcome (locomotor skills) data [19]. Procedures and methods in both studies conform to ethical guidelines defined by the World Medical Association's Declaration of Helsinki and its subsequent revisions. . The Norwegian South-East Regional Committee for Medical Research Ethics and the Norwegian Centre for Research Data approved the study protocols. We obtained written informed consent from each child's parents or legal guardians and from the responsible preschool and school authorities prior to all testing.

Procedures

Physical activity

PA was measured using the ActiGraph GT3X+ accelerometer (Pensacola, FL, USA) [20] worn at the waist over seven (ASK) and 14 (PRESPAS) consecutive days, except during water activities (swimming, showering) or while sleeping. Units were initialized at a sampling rate of 30 Hz and files were analyzed restricted to hours 06:00 to 23:59 using 1-second epochs to capture low and high intensity PA [21] using the KineSoft analytical software version 3.3.80 (KineSoft, Loughborough, UK). We applied wear time requirements of ≥ 8 hours/day and ≥ 4 days/week to constitute a valid measurement [22, 23].

As described previously, we applied “spectrum descriptions” of PA to obtain a better and more nuanced picture of associations between PA and the outcomes compared to traditional blunt descriptions of data. We have previously used descriptions of 23 variables (from 0–99, 100–249, 500–999, 1000–1499, … 9500–9999, to ≥ 10000 cpm) in the ASK dataset [3] and 33 variables (from 0–99, 100–249, 500–999, 1000–1499, … 14500–14999, to ≥ 15000 cpm) in the PRESPAS dataset [17], obtained from the vertical axis, to capture movement in narrow intensity intervals across the intensity spectrum. Thus, except for the two lowest bins and the highest bin, intervals of 500 cpm were used in the original studies. For simplicity, we created new bins with intervals of 1000 cpm for both datasets herein. Thus, the ASK data was described using 12 bins (from 0–99, 100–999, 1000–1999, 2000–2999, … 9000–9999, to ≥ 10000 cpm) and the PRESPAS data was described using 17 bins (from 0–99, 100–999, 1000–1999, 2000–2999, … 14000–14999, to ≥ 15000 cpm). These descriptions performed similarly to previously published models using higher resolution (explained variances = 17.02% for 500 cpm intervals and 15.69% for 1000 cpm intervals in the ASK dataset and 6.70% for 500 cpm intervals and 6.84% for 1000 cpm intervals in the PRESPAS dataset). In addition, we used the original descriptions of 23 and 33 bins in the ASK and PRESPAS datasets, respectively, as a basis for examining the sensitivity of the analyses to different resolution of the data across the intensity spectrum. Specifically, in each dataset, we constructed descriptions that were condensed to 1) three bins *below* 4000 cpm (0–99, 100–1999, and 2000–3999), where the original bins above 4000 cpm were retained, and 2) three bins *above* 4000 cpm (4000–5999, 6000–7999, and ≥ 8000 cpm in the ASK dataset; 4000–7999, 8000–11999, and ≥ 12000 cpm in the PRESPAS dataset), where the original bins below 4000 cpm were retained.

We used the Evenson et al [24, 25] intensity cut points of 0–99, 100–2295, 2296–4011, and ≥ 4012 cpm as a guidance for interpreting intensities across the spectrum as SED, LPA, MPA, and VPA *post hoc*.

Outcomes

The outcome in the ASK study was a cardiometabolic composite score [26] calculated as the mean of six variables (systolic blood pressure, triglycerides, total:high-density lipoprotein cholesterol, homeostasis model assessment of insulin resistance, waist:height ratio, and the inverse Andersen aerobic fitness test) using standardized scores after adjustment for sex and age using residuals from linear regression. A higher score indicates poorer cardiometabolic health.

The outcome in the PRESPAS study was a sum score of three locomotor movement tasks (run, horizontal jump, hop) guided by the Test of Gross Motor Development 3 test battery [27, 28]. A higher score indicates better locomotor skills. Children were scored quantitatively based on whether they did or did not demonstrate specific criteria for each skill based on the original scoring procedures. The criteria scores were averaged for each task and the total locomotor score. The score was standardized after adjustment for sex, age, body mass index, preschool, and assessor of motor skills using residuals from linear regression prior to analysis.

Statistical analyses

We analyzed PA profiles as raw data (i.e., min/day), normalized data (proportion of wear time), log-transformed data, and as compositional data. Compositional transformation was performed using the centered log-ratio (clr) method as described by Hinkle and Rayens [29] and by constructing log-ratios according to SED (i.e., 0–99 cpm, SED log-ratio (SEDlr)), LPA (i.e., 1000–1999 cpm, LPAlr), MPA (i.e., 3000–3999 cpm, MPAlr), and VPA (i.e., the variable strongest associated with the outcome using absolute data = 7000–7999 cpm in the ASK dataset and 5000–5999 cpm in the PRESPAS dataset, VPAlr). Because multivariate pattern analysis can handle any degree of multicollinearity, the use of the isometric log-ratio [30] was not necessary (the isometric log-ratio technically provides an open dataset and makes it possible to analyze data by linear regression in the absence of other sources of multicollinearity than closure, i.e., a correlation of -1 spread over the explanatory variables). Moreover, the isometric log-ratio transform does not allow for determination of the association pattern from one joint interpretable model (i.e., separate models has to be performed for each intensity bin). Centred log-ratios were made by 1) normalization of the PA intensity profile for each individual, that is, calculation of proportions where all explanatory variables sum to 1, prior to 2) making natural log-transformations of each variable, and 3) centering the explanatory variables for each object to the mean logarithm of all explanatory variables (clr) [31]. Construction of other log-ratios (SEDlr, LPAlr, MPAlr, VPAlr) followed steps 1 and 2, but used the selected PA bins (SED, LPA, MPA, and VPA) as the denominator.

Multivariate pattern analysis. Partial least squares (PLS) regression analysis [14] was used to determine the multivariate association patterns of raw and compositional PA data (explanatory variables) with cardiometabolic health (ASK dataset) and locomotor skills (PRESPAS dataset) (outcomes) (please see Aadland et al [3, 32] for a brief overview of this statistical approach as applied to PA accelerometry data). PLS regression decomposes the explanatory variables into orthogonal linear combinations (PLS components), while simultaneously maximizing the covariance with the outcome variable. Thus, PLS regression is able to handle completely collinear variables through the use of latent variable modelling [14]. The procedure differs from that of factor analysis or principal component analysis by creating components that maximize the covariation with the outcome, not internally among the explanatory variables. Prior to PLS regression, all variables were centered and standardized to unit variance. Models were cross-validated using Monte Carlo resampling with 1000 repetitions by repeatedly and randomly keeping 50% of the subjects as an external validation set when estimating the models to determine the number of PLS components [33]. Validation is an integrated part of the procedure to avoid overfitting due to inclusion of minor PLS components representing noise. For each validated PLS regression model, a single predictive component was subsequently calculated by means of target projection [15, 11] to express all the predictive variance in the PA intensity spectrum related to cardiometabolic health in a single intensity vector. Selectivity ratios (SRs) with 95% CIs were obtained as the ratio of this explained predictive variance to the total variance for each PA intensity variable [32, 34, 35]. Thus, the SRs provide the explained variance of each variable with the predicted outcome in the multivariate space, while retaining their direction of association. The procedure for obtaining the multivariate patterns is completely data-driven, with no assumptions on variable distributions or degree of collinearity among variables. These analyses were performed by means of the commercial software Sirius version 11.0 (Pattern Recognition Systems AS, Bergen, Norway).

Geometric means were calculated for the mean logarithm of all explanatory variables to compare the center of different clr models. Pearson's correlation was used to compare association patterns for raw, normalized, and log-transformed data.

Results

We included 841 schoolchildren (mean (SD) 10.2 (0.3) years old, 50% boys) and 1081 preschool children (4.7 (0.9) years old, 52% boys) who provided valid data on all relevant variables.

Figures 1 and 2 show the multivariate association patterns between PA and cardiometabolic health (ASK dataset) and between PA and locomotor skills (PRESPAS dataset), respectively, using raw (min/day), normalized (proportion of wear time), log-transformed, and compositional data. Within both datasets, explained variances were similar for raw and normalized data, but improved for log-transformed and compositional data. Still, association patterns were similar for models using absolute (i.e., raw, normalized, and log-transformed) data (ASK: $r \geq 0.99$; PRESPAS: $r \geq 0.93$). In the ASK dataset, the strongest association with cardiometabolic health was found for 7000–7999 cpm when using absolute data. Consistent with this finding, the compositional data also showed reallocating time from lower intensities to 7000–7999 cpm was most favorable. In contrast, while 5000–5999 cpm was strongest associated with locomotor skills in the PRESPAS dataset when using absolute data, compositional data showed reallocating time from lower intensities to 12000–12999 cpm was most favorable. The centers (geometric means) of the compositional data were 0.0163 and 0.0070 in the ASK and PRESPAS datasets (Supplemental Figure 1), respectively, having corresponding turning points for favorable/unfavorable reallocation of time around 4000–4999 and 7000–7999 cpm (Figures 1 and 2).

To further explore the possible impact of including higher intensities and having a higher number of variables in the PRESPAS dataset (0–99 to ≥ 15000 cpm; 17 variables) than in the ASK dataset (0–99 to ≥ 10000 cpm; 12 variables), we tested association patterns using datasets with different resolution of data across the intensity spectrum. In both the ASK dataset (Figure 3) and the PRESPAS dataset (Figure 4), varying the resolution had a clear impact on geometric means and association patterns using compositional data. In the ASK dataset, the turning points (for which time should be reallocated around) were 5500–5999 (geometric mean = 0.0060) and 3000–3499 cpm (geometric mean = 0.0249) when using datasets that were condensed below and above 4000 cpm, respectively. This difference was more pronounced in the PRESPAS dataset, for which corresponding turning points of 9500–9999 (geometric mean = 0.0023) and 2500–2999 cpm (geometric mean = 0.0270) were found.

In addition to using the data-driven centered log-ratio (which use the geometric mean composition as its reference for constructing log-ratios), we explored how the use of alternative log-ratios (SEDlr, LPAlr, MPAlr, and VPAlr) performed. In the ASK dataset (Figure 5), reallocating time from SED (0–99 cpm) and LPA (2000–2999 cpm) to all other intensities were favorable, whereas the findings for MPA (3000–3999 cpm) were mixed (i.e., favorable to reallocate time from lower intensities to MPA and unfavorable to reallocate time from higher intensities to MPA). Reallocating time from VPA (i.e., the intensity strongest associated with cardiometabolic health using absolute data; 7000–7999 cpm) to all other intensities were unfavorable. In the PRESPAS dataset (Figure 6), the association patterns were rather similar to those found in the ASK dataset. However, when using VPAlr (i.e., the intensity strongest associated with locomotor skills using absolute data (5000–5999 cpm) as the denominator for constructing compositions), the results showed reallocating time from 5000–5999 cpm to higher intensities was favorable, and thus similar to findings for clr, but in conflict with the interpretation using absolute data.

Discussion

In the present study we used two large datasets including children of different age to explore association patterns between two different outcomes and spectrum descriptions of PA using raw, normalized, log-transformed, and compositional data. Our findings provide a good basis for elaboration on interpretation of different models, and show some similarities, but also highlight some challenges and dissimilar interpretations. Briefly, while all findings show that intensities in the vigorous range is strongest associated with the outcomes in both datasets, the relative nature of the compositional data leads to some counterintuitive findings and differences in interpretation caused by the operationalisation of PA data. These findings challenge and confuse interpretation of association patterns derived from the different approaches. Our findings therefore suggest future studies using accelerometry-derived PA data apply multivariate pattern analysis with raw data, or possibly log-transformed data, to avoid confusion in developing the field of PA epidemiology.

While associations for raw, normalized, and log-transformed data are *absolute* (i.e., associations are given as each variable's importance for the predicted outcome in a multivariate space), associations for compositional data are *relative* (i.e., associations for all variables are relative to other variables). Thus, although the association patterns in Figures 1 and 2 appear different, the interpretation of the patterns are partly equivalent. For the ASK dataset (Figure 1), the use of absolute data shows vigorous intensities are strongly associated with cardiometabolic health, whereas associations for lower intensities are weak. For comparison, the compositional data shows that reallocating time from lower intensities to vigorous intensity is favourable. This interpretation can also be made from the absolute data; since spending more time in VPA is favourable, spending less time in VPA – and thus more time in lower intensities – is unfavourable. However, the results from the compositional data presents three major challenges for interpretation.

First, since associations are relative, it is not possible to determine how each variable is associated with the outcome. This challenge also, to some extent, applies to interpretation of isotemporal substitution models. For example, if it is favourable to reallocate time from SED to VPA, it will by definition also be unfavourable to reallocate time from VPA to SED, but does that mean spending more time in VPA is favourable or more time in SED in unfavourable (or both)? The compositional data in Figure 1 clearly suggests spending time in SED, LPA, and MPA are as detrimental to cardiometabolic health as VPA is favourable, because the associations are mirrored around a turning point, that is, the centre of the distribution (the geometric mean composition). In essence, this approach reduces findings to the trivial interpretation that one should substitute all PA activities with an intensity below the centre point to activities with an intensity above this point, because it is not possible to determine the intensities' absolute relation to the outcome. This point is also underpinned by constructing alternative log-ratios (Figures 5 and 6), of which nearly all provide the hypothesized association patterns. For example, while the SEDlr shows reallocating time from SED to higher intensities is favourable, the VPAlr shows reallocating time from VPA to lower intensities is unfavourable. Still, similar to the clr, the VPAlr would easily be misinterpreted and taken as evidence of an unfavourable association between low intensities (up to 6999 cpm) and cardiometabolic health, though this association pattern is relative and trivial.

Second, all findings using compositional data suggest it is more unfavourable to substitute time in higher intensities with time in LPA than in SED. This finding contradict current evidence indicating a possible favourable association between LPA and cardiometabolic health [6] but no association for SED and cardiometabolic health [36], and our raw data showing a weak unfavourable association for SED, but no association for LPA. We simply suggest this finding is a statistical artefact resulting from the great alteration of the correlation structure of the explanatory data matrix [3]. This enforced correlation structure might also result in the major shift in the intensities that are most favourably associated with locomotor skills in the PRESPAS dataset (Figure 2). While the strongest association is found for 5000–5999 cpm using absolute data, the analysis of compositional data suggest it is most favourable to spend time in 12000–12999 cpm. We suggest these differing findings result from the nature of the compositional transformation; since associations are relative, associations are not favourable or unfavourable, but rather more favourable or less favourable, which leads to a biphasic association pattern around the geometric mean composition. This turning point will inherently be in the middle of the distribution, which is where we (accidentally) find the strongest absolute association in the PRESPAS dataset. A similar pattern is seen for the VPAlr (Figure 6); since the variable showing the strongest association with locomotor skills was used for constructing the log-ratio, we would expect reallocating time from this intensity to all others intensities would be unfavourable. In contrast, reallocating time to higher intensities seems favourable. These findings result from the enforced biphasic association pattern, driven by the nature of the log-ratio. Interestingly, the geometric mean composition essentially provide the same information as the intensity gradient, introduced by Rowlands et al [37], as both measures capture and condense the distribution of time spent across the intensity spectrum to a single variable (r between intensity gradient and geometric mean = 0.96 in the ASK dataset and 0.97 in the PRESPAS dataset, results not reported). A higher geometric mean means more time is spent in higher intensities, which means the slope across the intensity spectrum is less steep. Thus, a higher geometric mean means the intensity distribution is moved upwards, resulting in lower intensities being negatively related to the mean and higher intensities being positively related to the mean using compositional data. This biphasic pattern, suggesting it is favourable to reallocate time from lower to higher intensities, is also seen when restricting the spectrum in PRESPAS to intensities above 5000 cpm (results not shown), for which weaker associations are seen for increasing intensities using absolute data (Figure 2). Thus, these approaches result in directly conflicting findings.

Third, another challenge for interpretation of compositional data is the major influence on findings resulting from different operationalisation or binning of PA data. As shown in Figures 3 and 4, different binning – of the same data – result in largely different interpretations. We hypothesized the clr would be vulnerable to the operationalisation of data, because the number of variables and their compositions would affect the geometric mean and thus the centre of the distribution. While the influence was moderate in the ASK dataset (the centre was found for 3000–3499 cpm versus 5500–5999 cpm for the datasets condensed above and below 4000 cpm, respectively), a major influence was found in the PRESPAS dataset (the centre was found for 2500–2999 cpm versus 9500–9999 cpm for the datasets condensed above and below 4000 cpm, respectively) for which a higher number of higher intensity variables was included. Thus, more or less arbitrary choices in variable selection or binning may lead to substantial differences in interpretation, which may confuse conclusions and hamper comparability between studies.

An assumption underlying the possible need for a compositional transformation of data, which, to the best of our knowledge, has not been discussed previously, is how PA variables across the intensity spectrum are associated with the outcome. The general idea behind constructing a closed dataset is that time needs to be reallocated among activities or intensities, that is, time cannot be added to some intensities without detracting from other intensities. However, most of the time over a day is spent sedentary or in low intensity, which is not important to health and developmental outcomes in children [9, 3, 17, 36]. Thus, there is arguably abundant time that that can be added to the higher intensities without influencing the association model. In the included datasets, 719 out of 795 minutes/day (ASK) and 627 out of 702 minutes/day (PRESPAS) (89–90% of the time) is accumulated in SED and LPA [3, 17], for which associations with the outcomes are very weak. Thus, it does not matter if some minutes are taken from the lower intensities to increase time in higher intensities. This point challenges the foundational assumption of compositional analysis and support the use of absolute data for analysis.

We have previously shown a great improvement of association models between PA and cardiometabolic health when using a more detailed PA description compared to traditional descriptions [9, 10]. This improved resolution and higher number of variables have several important implications for analysis. First, if using a blunt description of only two variables, for example SED and non-SED PA, given a constant sum of these variables (i.e., total wear time), they will be perfectly negatively correlated and thus singular (the imposed correlation of the explanatory data matrix is -1). In this situation, spending more time in one variable will lead to an equivalent reduction in time spent in the other, so time will be fully reallocated across these variables. In situations with few variables, this reallocation might be a concern. However, when analyzing an explanatory data matrix comprising a higher number of variables, reallocation of time will be spread over many variables and thus be less important when estimating associations [3, 32]. This feature makes the transformation to compositional data less relevant in this situation than for traditional descriptions of data (for example SED, LPA, and MVPA, or different types of movement). Also with few variables, though, whether raw or compositional, data will be correlated, which leaves linear regression less suited for analysis [3, 5]. Importantly, if using multivariate pattern analysis, the construction of compositions is not needed, since the model can handle singular data. Interestingly, though, log-transformation of data (centered or not) lead to improved model fit in both datasets, as also demonstrated for compositional data using both a traditional and spectrum description previously [3]. This finding suggests log-transformation reduces noise in the accelerometer data. It is well-known that count data produces heteroscedastic noise, which means that noise increases with the number of counts [16]. This source of noise is largest for SED and LPA. At the other end of the intensity spectrum, the number of counts is much lower, but the distributions are typically positively skewed. Skewed data may lead to a problem for modelling since validation and optimization of model selection (i.e., the number of PLS components included) is based on repeated Monte-Carlo resampling. The procedure use half of the sample for modelling and half of the sample for prediction, randomly partitioned for each repetition. Skewed distributions at the higher end of the PA intensity spectrum means that several PLS components that are weakly associated with the predicted outcome are needed to accommodate this variation between participants. The use of log-transformed data makes the distributions for these higher PA intensities less skewed, and thus more stable to resampling, which ultimately leads to simpler and more robust descriptions of data. This effect was most evident in the PRESPAS dataset for which we used the most detailed description of the highest intensities (up to ≥ 15000 cpm) (Figure 2).

Strengths and limitations

The main strength of the present study is the direct comparison of different analytic approaches to analyze associations between PA and two different outcomes in two large datasets. The use of these different datasets allowed for robust comparisons across the statistical approaches, and provided a nuanced picture of the findings beyond what would be possible with only one dataset. Importantly, we included two spectrum descriptions of PA, which, compared to traditional blunt descriptions, provide a better opportunity to reveal how the different handling of data affect the interpretation. Importantly, the same challenges as revealed herein apply to fewer variables, although less apparent.

Nevertheless, while a spectrum description may provide a substantially improved picture of association patterns than traditional description with fewer variables [9, 10], a simpler description might be needed for translation of findings into PA guidelines [32].

The cross-sectional designs limits our ability to draw conclusions about causality. It should also be kept in mind that use of other cohorts, for example spanning other age groups, and the use of other outcomes, could lead to other findings due to different correlation structures among the explanatory PA variables and/or different association patterns between PA intensities and outcomes. Further studies are therefore warranted to explore these analytic issues and extend our findings. Finally, we do not know the "true" association pattern between PA intensities and cardiometabolic health and motor skills. Thus, our conclusions of which approach provide the best results need to be interpreted with this issue in mind.

Conclusion

Interpretation of associations between accelerometer-derived PA spectra and cardiometabolic health and motor skills in two different samples of children differed when analyzing absolute (raw, normalized, and log-transformed) and relative (compositional) data. While we find interpretation of association patterns for absolute data meaningful, we find interpretation of association patterns for compositional data challenging and partly in conflict with results from absolute data. Moreover, our findings show that the relative nature of compositional data makes interpretation of association patterns susceptible to change according to how the explanatory data matrix is operationalized. Consistent with a previous study [3], we therefore recommend future studies using accelerometry-derived PA data apply multivariate pattern analysis with raw data, or possibly log-transformed data, to avoid confusion in developing the field of PA epidemiology. Finally, we find the use of absolute data more meaningful for development and messaging of PA guidelines, as results can be interpreted directly according to the strength of association between the PA variables and the outcome, and directly in minutes of PA spent in given intensities of activity, although this is complicated by the correlated data structure [32].

Funding

The ASK study was funded by the Research Council of Norway (grant number 221047/F40) and the Gjensidige Foundation (grant number 1042294). The PRESPAS study was funded by the Sogn og Fjordane County Municipality. Jairo H. Migueles is supported by a Grant from the Spanish Ministry of Education, Culture and Sport (FPU15/02645). Additional funding was obtained from the SmarterMove project supported by the MINECO/FEDER (DEP2016-79512-R) and the University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence; Scientific Excellence Unit on Exercise and Health (UCEES). Junta de Andalucía, Consejería de Conocimiento, Investigación y Universidades and European Regional Development Funds (ref. SOMM17/6107/UGR). None of the funding agencies had any role in the study design, data collection, analyzing or interpreting data, or in writing the manuscripts.

Acknowledgements

We thank all children, parents and staff at the participating preschools (PRESPAS) and schools (ASK) for their excellent cooperation during the data collection. We also thank colleagues and students at the *Western Norway University of Applied Sciences* (formerly *Sogn og Fjordane University College*) for their contribution to the ASK and PRESPAS studies. We thank participants at the *International Workshop: A focus on statistical methods to analyse accelerometer-measured physical activity*, Granada, Spain, October 21.-22. 2019 for their valuable perspectives on best practices for analyzing physical activity data.

Availability of data and materials

The datasets used in the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

References

1. Chastin SFM, Palarea-Albaladejo J, Dontje ML, Skelton DA. Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. Plos One. 2015;10(10). doi:10.1371/journal.pone.0139984.

2. Dumuid D, Stanford TE, Martin-Fernandez JA, Pedisic Z, Maher CA, Lewis LK et al. Compositional data analysis for physical activity, sedentary time and sleep research. Stat Methods Med Res. 2018;27(12):3726-38. doi:10.1177/0962280217710835.

3. Aadland E, Kvalheim OM, Anderssen SA, Resaland GK, Andersen LB. Multicollinear physical activity accelerometry data and associations to cardiometabolic health: challenges, pitfalls, and potential solutions. Int J Behav Nutr Phys Act. 2019;16(1). doi:10.1186/s12966-019-0836-z.

4. Mekary RA, Willett WC, Hu FB, Ding EL. Isotemporal Substitution Paradigm for Physical Activity Epidemiology and Weight Change. Am J Epidemiol. 2009;170(4):519-27. doi:10.1093/aje/kwp163.

5. Cohen J, Cohen P, West SG, Aiken LS. Applied multiple regression/correlation analysis for the bahavioral sciences. 3 ed. New York: Routledge; 2003.

6. Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput JP, Janssen I et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. Appl Physiol Nutr Metabol. 2016;41(6):S197-S239. doi:10.1139/apnm-2015-0663.

7. van der Ploeg HP, Hillsdon M. Is sedentary behaviour just physical inactivity by another name? Int J Behav Nutr Phys Act. 2017;14:8. doi:10.1186/s12966-017-0601-0.

8. Pedisic Z. Measurement issues and poor adjustments for physical activity and sleep undermine sedentary behaviour research - the focus should shift to the balance between sleep, sedentary behaviour, standing and activity. Kinesiology. 2014;46(1):135-46.
9. Aadland E, Kvalheim OM, Anderssen SA, Resaland GK, Andersen LB. The multivariate physical activity signature associated with metabolic health in children. Int J Behav Nutr Phys Act. 2018;15(77). doi:10.1186/s12966-018-0707-z.
10. Aadland E, Kvalheim OM, Anderssen SA, Resaland GK, Andersen LB. The Triaxial Physical Activity Signature Associated with Metabolic Health in Children. Med Sci Sports Exerc. 2019;51(10):2173-9. doi:10.1249/mss.0000000000002021.
11. Rajalahti T, Kvalheim OM. Multivariate data analysis in pharmaceutics: A tutorial review. Int J Pharm. 2011;417(1-2):280-90. doi:10.1016/j.ijpharm.2011.02.019.
12. Rajalahti T, Kroksveen AC, Arneberg R, Berven FS, Vedeler CA, Myhr K-M et al. A multivariate approach to reveal biomarker signatures for disease classification: application to mass spectral profiles of cerebrospinal fluid from patients with multiple sclerosis. J Proteome Res. 2010;9(7):3608-20. doi:10.1021/pr100142m.
13. Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics-A review in human disease diagnosis. Anal Chim Acta. 2010;659(1-2):23-33. doi:10.1016/j.aca.2009.11.042.
14. Wold S, Ruhe A, Wold H, Dunn WJ. The collinearity problem in linear-regression - the partial least-squares (pls) approach to generalized inverses. SIAM J Sci Comput. 1984;5(3):735-43. doi:10.1137/0905052.
15. Kvalheim OM, Karstang TV. Interpretation of latent-variable regression-models. Chemometr Intell Lab Syst. 1989;7(1-2):39-51. doi:10.1016/0169-7439(89)80110-8.
16. Kvalheim OM, Brakstad F, Liang Y. Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. Anal Chem. 1994;66(43-51).
17. Nilsen AKO, Anderssen SA, Loftesnes JM, Johannessen K, Ylvisaaker E, Aadland E. The multivariate physical activity signature associated with fundamental motor skills in preschoolers. J Sports Sci. 2019:1-9. doi:10.1080/02640414.2019.1694128.
18. Resaland GK, Moe VF, Aadland E, Steene-Johannessen J, Glosvik Ø, Andersen JR et al. Active Smarter Kids (ASK): Rationale and design of a cluster-randomized controlled trial investigating the effects of daily physical activity on children's academic performance and risk factors for non-communicable diseases. BMC Public Health. 2015;15:709-. doi:10.1186/s12889-015-2049-y.
19. Nilsen AKO, Anderssen SA, Resaland GK, Johannessen K, Ylvisaaker E, Aadland E. Boys, older children, and highly active children benefit most from the preschool arena regarding moderate-to-vigorous physical activity: A cross-sectional study of Norwegian preschoolers. Prev Med Reports. 2019;14:100837-. doi:10.1016/j.pmedr.2019.100837.

20. John D, Freedson P. ActiGraph and Actical physical activity monitors: a peek under the hood. Med Sci Sports Exerc. 2012;44(1 Suppl 1):S86-S9.

21. Aadland E, Andersen LB, Anderssen SA, Resaland GK, Kvalheim OM. Accelerometer epoch setting is decisive for associations between physical activity and metabolic health in children. J Sports Sci. 2019:1-8. doi:10.1080/02640414.2019.1693320.

22. Aadland E, Andersen LB, Skrede T, Ekelund U, Anderssen SA, Resaland GK. Reproducibility of objectively measured physical activity and sedentary time over two seasons in children; Comparing a day-by-day and a week-by-week approach. Plos One. 2017;12(12). doi:10.1371/journal.pone.0189304.

23. Aadland E, Johannessen K. Agreement of objectively measured physical activity and sedentary time in preschool children. Prev Med Reports. 2015;2:635-9.

24. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008;26(14):1557-65. doi:10.1080/02640410802334196.

25. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of Accelerometer Cut Points for Predicting Activity Intensity in Youth. Med Sci Sports Exerc. 2011;43(7):1360-8. doi:10.1249/MSS.0b013e318206476e.

26. Andersen LB, Harro M, Sardinha LB, Froberg K, Ekelund U, Brage S et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). Lancet. 2006;368(9532):299-304. doi:10.1016/S0140-6736(06)69075-2.

27. Ulrich D. Test of Gross Motor Development – Third Edition. Ann Arbor, MI: Center on Physical Activity and Health in Pediatric Disabilities, 2013.

28. Ulrich DA. Test of Gross Motor Development - Third edition. Examiner's Manual. Austin, Texas: Pro.ed.; 2019.

29. Hinkle J, Rayens W. Partial least squares and compositional data: problems and alternatives. Chemometr Intell Lab Syst. 1995;30(1):159-72. doi:https://doi.org/10.1016/0169-7439(95)00062-3.

30. Hron K, Filzmoser P, Thompson K. Linear regression with compositional explanatory variables. J Appl Stat. 2012;39(5):1115-28. doi:10.1080/02664763.2011.644268.

31. Aitchison J. The statistial analysis of compositional data. Journal of the Royal Statistical Society,. 1982;44(2):139-77.

32. Aadland E, Andersen LB, Resaland GK, Kvalheim OM. Interpretation of Multivariate Association Patterns between Multicollinear Physical Activity Accelerometry Data and Cardiometabolic Health in Children—A Tutorial. Metabolites. 2019;9(7):129.

33. Kvalheim OM, Arneberg R, Grung B, Rajalahti T. Determination of optimum number of components in partial least squares regression from distributions of the root-mean-squared error obtained by Monte Carlo resampling. J Chemometrics. 2018. doi:10.1002/cem.2993.

34. Rajalahti T, Arneberg R, Berven FS, Myhr KM, Ulvik RJ, Kvalheim OM. Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. Chemometr Intell Lab Syst. 2009;95(1):35-48. doi:10.1016/j.chemolab.2008.08.004.

35. Rajalahti T, Arneberg R, Kroksveen AC, Berle M, Myhr KM, Kvalheim OM. Discriminating Variable Test and Selectivity Ratio Plot: Quantitative Tools for Interpretation and Variable (Biomarker) Selection in Complex Spectral or Chromatographic Profiles. Anal Chem. 2009;81(7):2581-90. doi:10.1021/ac802514y.

36. Cliff DP, Hesketh KD, Vella SA, Hinkley T, Tsiros MD, Ridgers ND et al. Objectively measured sedentary behaviour and health and development in children and adolescents: systematic review and meta-analysis. Obes Rev. 2016;17(4):330-44. doi:10.1111/obr.12371.

37. Rowlands AV, Edwardson CL, Davies MJ, Khunti K, Harrington DM, Yates T. Beyond Cut Points: Accelerometer Metrics that Capture the Physical Activity Profile. Med Sci Sports Exerc. 2018;50(6):1323-32. doi:10.1249/MSS.0000000000001561.

Figure Legends

Figure 1. Association patterns between physical activity intensities and a composite cardiometabolic health score (ASK dataset) using raw data (min/day), normalized data (proportions of wear time), log-transformed data, and compositional data (clr). The models included 4, 4, 4, and 3 PLS components, respectively. Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable

Figure 2. Association patterns between physical activity intensities and locomotor skills (PRESPAS dataset) using raw data (min/day), normalized data (proportions of wear time), log-transformed data, and compositional data (clr). All models included 3 PLS components. Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable.

Figure 3. Association patterns between physical activity intensities and a composite cardiometabolic health score (ASK dataset) using two different descriptions of physical activity across the intensity spectrum. The data (originally described using 23 bins, from 0–99, 100–249, 500–999, 1000–1499, … 9500–9999, to ≥ 10000 cpm, [3]) were condensed to three bins below 4000 cpm (left panel) and three bins above 4000 cpm (right panel). Raw data (upper panel), compositional data (clr) (lower panel). Explained variances were 14.34% (3 components) for raw data and 18.37% (3 components) for compositional data when the data was condensed below 4000 cpm, and 16.98% (5 components) for raw data and 20.76% (5 components) for compositional data when the data was condensed above 4000 cpm. Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable.

Figure 4. Association patterns between physical activity intensities and locomotor skills (PRESPAS dataset) using two different descriptions of physical activity across the intensity spectrum. The data (originally described using 33 bins, from 0–99, 100–249, 500–999, 1000–1499, … 14500–14999, to ≥ 15000 cpm, [17]) were condensed to three bins below 4000 cpm (left panel) and three bins above 4000 cpm (right panel). Raw data (upper panel), compositional data (clr) (lower panel). Explained variances were 5.67% (1 components) for raw data and 7.99% (3 components) for compositional data when the data was condensed below 4000 cpm, and 6.35% (2 components) for raw data and 7.96% (2 components) for compositional data when the data was condensed above 4000 cpm. Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable.

Figure 5. **Association patterns between physical activity intensities and a composite cardiometabolic health score (ASK dataset) using four different log-ratios of physical activity across the intensity spectrum.** SEDlr: reference = 0–99 cpm, explained variance = 20.01% (5 components) (upper left panel); LPAlr: reference = 1000–1999 cpm, explained variance = 20.75% (4 components) (upper right panel); MPAlr: reference = 3000–3999 cpm, explained variance = 18.49% (2 components) (lower left panel); VPAlr: reference = 7000–7999 cpm (i.e., the intensity strongest associated with cardiometabolic health using raw data), explained variance = 19.90% (5 components) (lower right panel). Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable.

Figure 6. **Association patterns between physical activity intensities and locomotor skills (PRESPAS dataset) using four different log-ratios of physical activity across the intensity spectrum.** SEDlr: reference = 0–99 cpm, explained variance = 7.94% (3 components) (upper left panel); LPAlr: reference = 1000–1999 cpm, explained variance = 7.60% (1 components) (upper right panel); MPAlr:

reference = 3000–3999 cpm, explained variance = 8.37% (2 components) (lower left panel); VPAlr: reference = 5000–5999 cpm (i.e., the intensity strongest associated with cardiometabolic health using raw data), explained variance = 8.09% (3 components) (lower right panel). Selectivity ratios are calculated as explained to total variance on the predictive (target projected) component, while retaining the direction of association for each variable.

Supplemental Figure 1. Proportions of wear time spent in different PA intensities in the ASK dataset (left panel) and PRESPAS dataset (right panel). Horizontal lines shows the geometric mean composition.