Received: 29 April 2021 | Revised: 1 September 2021 | Accepted: 8 September 2021

DOI: 10.1111/sms.14052

ORIGINAL ARTICLE

WILEY

Physical activity spectrum discriminant analysis—A method to compare detailed patterns between groups

Jonatan Fridolfsson[1] | **Daniel Arvidsson**[1] | **Lars Bo Andersen**[2] | **Ola Thorsson**[3,4] | **Per Wollmer**[3,4] | **Björn Rosengren**[5] | **Magnus K. Karlsson**[5] | **Magnus Dencker**[3,4]

[1]Department of Food and Nutrition and Sports Science, Center for Health and Performance, University of Gothenburg, Gothenburg, Sweden

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

[3]Department of Translational Medicine, Lund University, Malmö, Sweden

[4]Department of Medical Imaging and Physiology, Skåne University Hospital, Malmö, Sweden

[5]Clinical and Molecular Osteoporosis Research Unit, Department of Clinical Sciences and Orthopaedics, Lund University, Skåne University Hospital, Malmö, Sweden

Correspondence

Jonatan Fridolfsson, Department of Food and Nutrition and Sport Science, Center for Health and Performance, University of Gothenburg, Box 300, 405 30 Gothenburg, Sweden.
Email: jonatan.fridolfsson@gu.se

Funding information

Skånes universitetssjukhus, Grant/Award Number: 98913; Lunds Universitet, Grant/Award Number: 46118; Region Skåne, Grant/Award Number: 819961

Investigating physical activity (PA) patterns as a detailed intensity spectrum instead of crude intensity categories have improved the ability to analyze the relationship between measured PA and health variables. The aim of this methodological study was to introduce and investigate the utility of using detailed PA intensity spectrum compared to crude PA intensity categories for comparison of PA between groups and between repeated measures. The study sample consisted of two groups of children, where one group was scheduled for extended physical education (PE) by daily classes while the other group followed usual PE schedule. Accelerometer data was processed into traditional crude PA intensity categories and into detailed PA intensity spectrum. Multivariate partial least squares regression for discriminant analysis (PLS-DA) was applied for PA intensity spectrum group comparison and compared to traditional univariate statistical analysis. Repeated measures were investigated using independent PLS-DA as well as multilevel PLS-DA for paired analysis. While traditional analysis of crude PA intensity categories was unable to find any group differences, multivariate analysis of the PA intensity spectrum identified statistically significant differences. By the extension of multilevel PLS-DA for paired comparison, a clear difference in the PA intensity spectrum was demonstrated between repeated measures. In conclusion, analysis of detailed PA intensity spectrum demonstrates utility for comparing detailed PA data between groups and between repeated measures in interventional and observational research.

KEYWORDS

accelerometer, discriminant analysis, multicollinearity, multilevel, multivariate pattern analysis, partial least squares regression, statistics

1 | INTRODUCTION

Measurement of physical activity (PA) by accelerometry most often yields continuous PA intensity as output.[1] However, common practice is to reduce the information of this output into crude intensity categories by applying one to five intensity cut points.[2] The time spent in one or more of these categories is then used as variables in further statistical analyses, e.g. the average number of minutes spent at moderate-to-vigorous PA.[1] Although, it is possible to divide the continuous PA intensity into a more detailed spectrum of categories with more than 20 cut points. The main problem with this approach is that the categories are highly collinear.[3] In general, the purpose of measuring PA is further to analyze the relationship to other variables, such as risk factors for disease, or comparing PA between groups. Traditional statistics such as multiple linear regression and group comparison by analysis of variance cannot handle variables that are highly collinear.

Multivariate statistical analysis by partial least squares (PLS) regression has recently been introduced in PA research to analyze the relationship between the PA intensity spectrum and health variables.[4,5] The method is used as a multivariate substitute to traditional regression analysis. PLS analysis has been extensively applied in chemometrics after it was introduced in the 1980s because of its ability to deal with collinear variables.[6] PLS decomposes the information in the predictor variables (independent) into a number of latent variables in a way that maximizes the covariation with the response variable (dependent). In the analysis of a PA spectrum, the latent variables represent different PA patterns that are related to the response variable, e.g. cardiometabolic health. By successively identifying PA patterns that are orthogonal to each other, collinearity is removed.

In addition to regression, PLS can also be used for analyzing group differences by PLS for discriminant analysis (PLS-DA).[7] By using a dummy variable to represent group belonging as a response variable in the PLS analysis, the pattern in the predictor variables that best discriminates the groups can be identified. By this way, PLS-DA can be used as a multivariate substitute for independent samples *t*-test for group comparison. PLS-DA can also be extended to multivariate PLS-DA to discriminate paired predictor variables.[8] By removing the between subject variation in the predictor variables, only the within subject variation is used as a predictor variable in order to find a pattern that discriminates each data pair. In this case, PLS-DA enables investigation of multivariate data in longitudinal studies with repeated measures. Consequently, PLS-DA would further advance the utility of the more detailed PA intensity spectrum in PA research.

The aim was to introduce and investigate the utility of using detailed PA intensity spectrum compared to crude PA intensity categories for comparison of PA between groups and between repeated measures.

2 | MATERIALS AND METHODS

2.1 | Study design

This is a methodological study investigating the application of a PA intensity spectrum together with PLS-DA. Previously collected data material with accelerometer measured PA in children was reanalyzed.[9,10] The original project investigated the health effect by implementing daily physical education (PE) compared to the ordinary schedule in primary school. Accelerometer data collection was first implemented into the project 2 years after the start of the added PE. Consequently, the repeated accelerometer data collection reflects a longitudinal follow-up rather than a strict intervention study with baseline. The first PA measurements were carried out during the fall in 2001 and in 2002 and were followed up by second measurements of the same individuals during the fall 2 years after the initial measurement.

In the present study, the PA intensity spectrum at the first measure (M1) was compared between two groups (between group analysis). Further, the PA intensity spectrum in one of the groups was compared between two repeated measures (M1 and M2) separated by 2 years in time (within group analysis). Traditional and multivariate analyses were applied and compared. In order to isolate the novel components of the statistical method and facilitate interpretation, the analyses in this study were set up as basic scenarios not considering school clustering or other more advanced multilevel approaches as well as other control variables. A future step would be to explore more complex models for expanded utilities.

2.2 | Study sample

Children from four schools in Malmö, Sweden, attending third or fourth grade at the time of the first measurement, were invited to participate in the original project. In one of the schools, children were scheduled for 40 min of PE daily during school hours, in total 200 min per week.[10] Children in the other three schools were scheduled for 60 min of PE per week, divided in 1–2 sessions, which is in line with the Swedish curriculum. The activities during the PE sessions followed the Swedish curriculum for PE and included a variety of ball games, running, jumping and climbing. The children from the school with daily PE will be referred to as the intervention group and the other children will be referred to as the control group. More

TABLE 1 Group characteristics

	Between group analysis		Within group analysis	
	Intervention M1	Control M1	Intervention M1	Intervention M2
N (% female)	128 (42%)	100 (46%)	93 (44%)	93 (44%)
Age (SD)	9.7 (0.6)	9.9 (0.6)	9.7 (0.6)	11.7 (0.6)

Note: The between group analysis included the intervention group and the control group, while the within group analysis only included the intervention group.

Abbreviations: M1, first measurement; M2, second measurement; N, number of participants with valid measurement; SD, standard deviation.

detailed descriptions of the recruitment of participants and the intervention content have been published previously.[9,10] Sample characteristics are shown in Table 1. The between school distribution of children in the control group was 25, 32 and 43.

The study was approved by the institutional ethics committee of Lund University (LU 243-01). Written informed consent was obtained by all participants' parents.

2.3 | PA measurement

Participants were instructed to wear a uniaxial accelerometer (MTI model 7164; Manufacturing Technology, Inc., also known as ActiGraph model 7164) for 4 days and removing it only when sleeping or during water activities. The accelerometer was worn on the right hip attached to an elastic band around the waist. Wear instructions were accompanied by written information, including a picture of correct sensor positioning. Vertical acceleration was recorded and output as ActiGraph counts with an epoch length of 10 s was generated.[1,11] Although a 10 s epoch length was used, intensity is referred to as counts per minute (CPM) in order to facilitate comparison to other studies. Non-wear time was defined as 60 min of consecutive zero output with allowance of up to 2 min of output equivalent to 100 CPM.[12] Accelerometer files where manually inspected to identify individuals who wore the accelerometer during sleep. In individuals where the accelerometer was worn during sleep, the output between 23 and 06 o'clock was set to zero. A valid day was defined as at least 8 h of wear time, and a valid measurement as at least three valid days. Two separate measurements of PA were performed with 2 years between, M1 and M2, both during the time of the intervention. The number of participants with valid PA measurement are shown in Table 1.

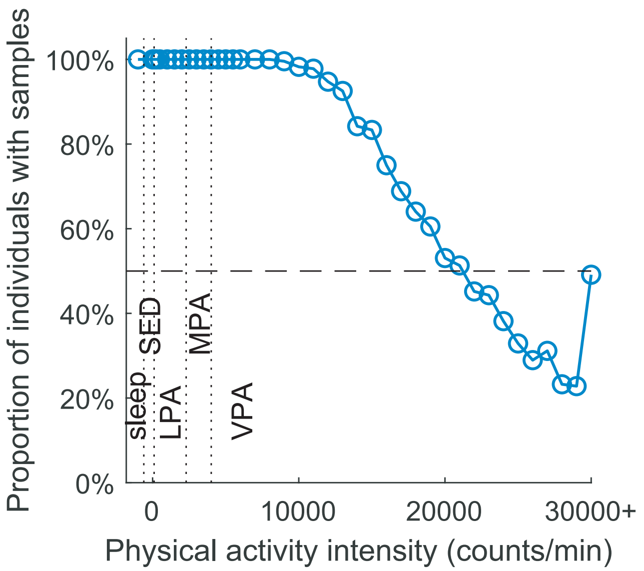

FIGURE 1 Proportion of individuals with at least one epoch at increasing physical activity (PA) intensity. Cut points for traditional PA intensity categories are presented (dotted vertical lines) as well as the PA intensity where only 50% of the participant provided accelerometer data (horizontal dashed line)

2.4 | Traditional statistical analyses

Traditional analysis of the PA intensity output was performed to be used as reference for the novel multivariate methods. Time spent at different crude intensity levels (sedentary [SED], light [LPA], moderate [MPA], vigorous [VPA] intensity) was identified by applying cut points developed by Evenson et al.[13] These cut points are the most frequently used in children.[2] In addition to the four intensity levels above, MPA and VPA were also jointed into a single variable representing the amount of PA commonly thought to be associated with health, usually referred to as moderate-to-vigorous intensity.[1] Independent samples *t*-tests were performed to examine the hypothesis that differences in time spent at the different PA intensities between the intervention and control groups were apparent.

2.5 | Multivariate statistical analyses

The PA intensity output was divided into a spectrum of small intervals (bins) up to the intensity where only 50%

of participants had at least one epoch (Figure 1). The rationale for the 50% threshold was that a previous study has shown that the relationship with health weakened with this proportion of individuals with data, likely caused by zero inflation.[5] At approximately 20 000 CPM 50% of participants had data. Intensity bins were created by setting cut points at 100, 250, 500, 1000, 1500, 2000..., 6000, 7000..., 19 000, 20 000, >20 000 CPM. This generated a PA intensity spectrum with 30 variables.

Comparison of PA intensity pattern between the intervention group and control group at M1 was undertaken by PLS-DA (between group analysis).[7] The 30 variables representing the PA intensity spectrum were standardized and used as predictor variables in the PLS analysis. The group belonging (intervention or control) was used as the response variable and was coded as a dummy variable, 1 representing the intervention group and −1 representing the control group. In this way, if the intervention group spent more time at a specific part of the PA intensity spectrum compared to the control group, there would be a positive association between the predictor variables corresponding to the specific intensity and the response variable. Consequently, the association would be negative if the control group was more physically active. PLS-DA decomposes the PA spectrum into several latent variables, referred to as PLS components, representing PA intensity patterns that best discriminate the two groups.

In addition, the PA intensity pattern was compared between M1 and M2 within the intervention group only (within group analysis), including individuals with valid measurements at both M1 and M2. Initially, PLS-DA was applied as above, considering the measurements as independent. The PA intensity spectrum variables were used as predictor variables and measurement occasion (M1 or M2) as response variable coded as −1 and 1 representing M1 and M2 respectively. PLS-DA finds the PA intensity patterns that best discriminate M1 and M2, treating the measurements as two independent groups, neglecting the underlying paired structure. An intensity specific increase of PA between M1 and M2 would result in a positive association between predictor and response variables, whereas a decrease result would result in a negative association.

In a second within group analysis of the difference in PA intensity patterns between M1 and M2, PLS-DA was used with a multilevel extension to consider the paired structure of the measurements. Similar to above, the response variable was coded as −1 and 1 representing M1 and M2 respectively. Between-subject variation was removed by considering the individual differences in the PA spectrum variables between M1 and M2 (M2 – M1). These differences were used as predictor variables to the response variable 1 representing M2 and the reversed difference (multiplied by −1) was used as predictor variables to the response variable −1 representing M1.[8] By including the difference twice in the same model, once by regressing it on to 1 and once by regressing it reversed on to −1, discrimination between the measurement occasions by PLS-DA is possible. PLS-DA finds the patterns of individual PA difference (from M1 to M2) that is associated with the discrimination of measurement occasion (from −1 to 1 representing M1 to M2). In this way, if the time spent at a specific part of the PA intensity spectrum increased from M1 to M2, there would be a positive association between the predictor variables corresponding to the specific intensity and the response variable. If it decreased, the association would be negative.

In PLS analysis, it is possible to generate as many PLS components (latent variables representing patterns) as there are predictor variables. However, this would overfit the model. Instead, the model requires cross-validation in order to find the optimal number of components.[6] Cross-validation was performed by Monte Carlo resampling with 1000 repetitions, randomly keeping 25% of the samples as validation samples. With an increasing number of components, the performance of the resampled PLS-DA models was evaluated by the proportion of correctly classified individuals.[14] Since the predictions of the PLS model are not limited to −1 and 1 but could take any continuous number, positive values were treated as 1 and negative as −1. The optimal number of components was selected by a backward selection procedure.[15] Initially, the number of components with the highest median proportion of correctly classified individuals was found. Thereafter, the proportion of resampled models with lower performance than the median performance of a model with one less component was calculated. This procedure was repeated with a decreasing number of components until this proportion was <0.401, corresponding to approximately 0.25 standard deviations, or only one component remaining.[15] By this, the number of components was selected to make sure that the performance of the model was substantially better than a model with fewer components.

Even with a high proportion of correctly classified individuals, the good performance may be attained by chance. To assess the uncertainty of the model, a permutation test was used.[16] Permutation tests compare the performance of the model to a null distribution of models generated by randomly permuting the order of the response variables values with regard to the predictor variables. The proportion of models generated from the permutated data with better performance than the original model represents a *p*-value indicating the statistical significance, that is the risk of making a type I error. In order to be able to attain a *p*-value of <0.01, 10^4 permutations were performed.[17]

The PA intensity pattern that the cross-validated model represents can be visualized using selectivity ratio plots, which identifies the PA intensity that is most influential for discriminating the groups.[18] The predictive performance of the PLS components was combined to a single component by target projection. The target-projected component was then used to calculate a selectivity ratio, which is the explained variance relative to the total variance of each predictor variable on the target-projected component.[18,19] Consequently, the explained variance of the selectivity ratio is standardized to the predictor variables' discriminative performance rather than the variation of the variables. Additionally, unstandardized PLS coefficients were visualized across the PA intensity spectrum to indicate absolute associations.[19] Furthermore, to show similarities between multivariate and univariate analysis, differences between group means of each variable on the PA intensity spectrum were also presented. This difference was presented both as standardized to each variable's standard deviation and as absolute difference in number of minutes per day. Confidence intervals of the selectivity ratio, PLS coefficients and univariate differences were calculated by bootstrapping with 10^4 repetitions.

Altogether, the outcome of the PLS-DA analyses is thus evaluated by the following measures: the strength of the PLS-DA models (discrimination performance) is determined by the proportion of correctly classified individuals; the statistical significance of the PLS-DA models is achieved from the permutation tests; finally, the PA intensity patterns that discriminate the groups or measurement occasions were combined by target projection and visualized by selectivity ratio plots. All data processing and statistical analyses were performed in MATLAB R2020a (MathWorks). PLS analysis was performed using the MATLAB function plsregress.

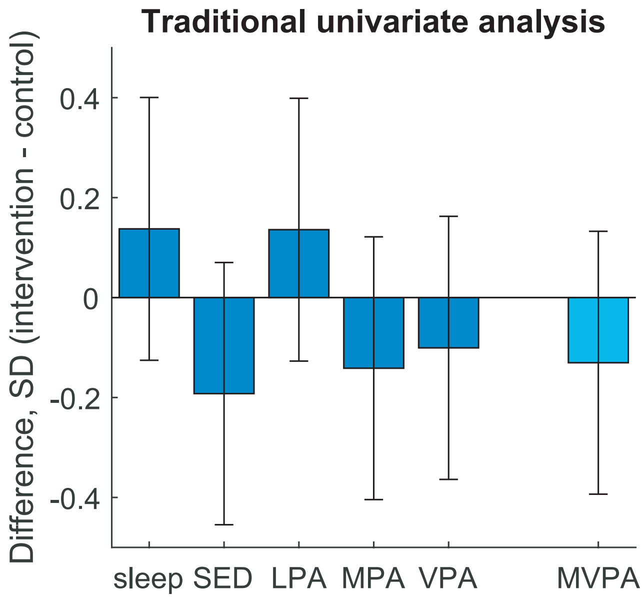

FIGURE 2 Group difference and 95% confidence interval in physical activity (PA) between intervention and control groups using traditional cut points. A positive difference indicates that the intervention group spent more time at this intensity level. SED, sedentary (0–103 counts per minute [CPM]); LPA, light PA (104–2295 CPM); MPA, moderate PA (2296–4011 CPM); VPA, vigorous PA (≥4012 CPM); MVPA, moderate-to-vigorous PA (MPA + VPA)

3 | RESULTS

We found no statistically significant differences in PA between the intervention and control groups using traditional crude intensity category cut points (Figure 2). When considering the entire PA intensity spectrum of the two groups using PLS-DA, the permutation test suggests a statistically significant group difference with $p < 0.01$. The intensity pattern was able to discriminate 72% of the individuals correctly. The multivariate pattern explaining the difference between groups is shown in Figure 3. Both the standardized and unstandardized patterns suggest that the intervention group performed less PA around the moderate-to-vigorous cut-point (3500–5000 CPM). In the standardized analysis (Figure 3A) it is apparent that the intervention group performed more very vigorous PA (above 10 000 CPM). This difference is not apparent in the unstandardized pattern (Figure 3B). On the other hand, the unstandardized pattern suggests the intervention group performed more LPA (750 CPM), which is not apparent in the standardized pattern. These multivariate patterns are similar to the univariate differences in group means across the PA intensity spectrum as shown in Figure 4. The collinearity between the crude intensity categories ranged from 16.6% to 67.7% whereas the collinearity between the intensity spectrum variables ranged from 2.5% to 94.2%.

The PA intensity pattern was different between M1 and M2 in the intervention group with p-values < 0.01 in both the independent group model and the paired model. However, the pattern of the independent model correctly discriminated 67% of the individuals whereas the pattern of the paired model correctly discriminated 87% of the individuals. The discriminating PA intensity patterns are shown as selectivity ratio plots in Figure 5. Both models show a decrease in LPA from M1 to M2. In addition, the paired model more clearly shows increased SED and decreased VPA between the measurements.

4 | DISCUSSION

The main result of this methodological study is that a difference in PA intensity pattern between the groups

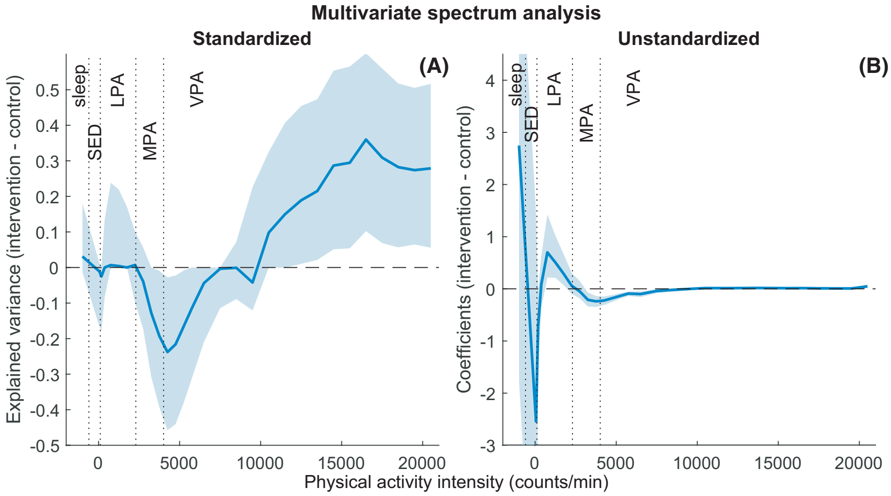

FIGURE 3 Multivariate pattern from partial lest squares regression for discriminant analysis (PLS-DA) to discriminate physical activity (PA) intensity spectrum between groups. (A) Standardized pattern represented by explained variance (selectivity ratio) and (B) unstandardized pattern represented by multivariate coefficients, both accompanied by 95% confidence interval. A positive value indicates that the intervention group spent more time at this intensity level. LPA, light PA; MPA, moderate PA; SED, sedentary; VPA, vigorous PA

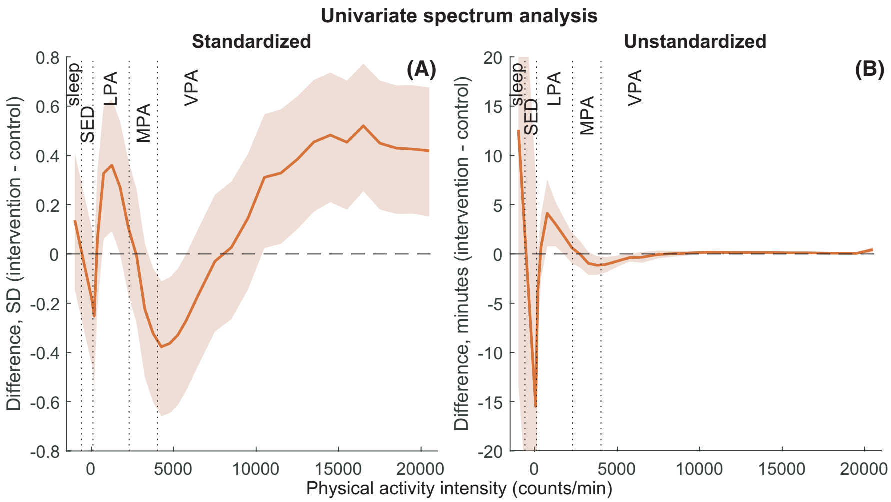

FIGURE 4 Group difference and 95% confidence interval across the physical activity (PA) intensity spectrum using univariate analysis to discriminate the PA intensity pattern. (A) *z*-Score standardized difference and (B) unstandardized difference in minutes per day. A positive difference indicates that the intervention group spent more time at this intensity level. LPA, light PA; MPA, moderate PA; SED, sedentary; VPA, vigorous PA

investigated was revealed with the detailed PA intensity spectrum only and not with the crude PA intensity categories. The intervention group performed less activity at the MPA to VPA transition while more activity at higher PA intensities. The PLS-DA provided a statistical measure of this difference, considering collinearity within the PA intensity spectrum. Further, the multilevel extension of PLS-DA generated a more robust model that clearly demonstrated the difference in PA intensity pattern between the repeated measures.

The group differences in PA intensity pattern identified by the analysis of the PA intensity spectrum could most likely have been found by traditional analysis if the group differences would have matched the crude cut points. In the present study, there was only a difference in the high part of the MPA range and not in the lower part. Furthermore, the control group performed more activity in the lower part of the VPA range whereas the intervention group performed more activity in the higher part of the VPA range. Therefore, this crude, and somewhat arbitrary, division of cut points is not able to identify the differences apparent. A previous study on the same sample has already shown a statistically significant group difference above 10 000 CPM.[20] They used different cut

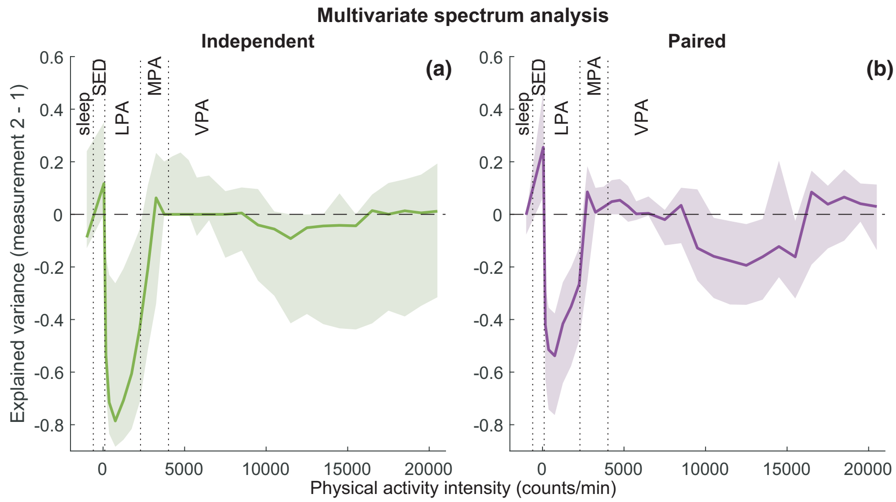

FIGURE 5 Explained variance (selectivity ratio) and 95% confidence interval in the physical activity (PA) intensity spectrum to discriminate the PA intensity pattern between the first and the second measurement of the intervention group. (A) Treating the measurements as independent. (B) Treating the measurements as paired. A positive value indicates that participants spent more time at this intensity level during the second measurement. LPA, light PA; MPA, moderate PA; SED, sedentary; VPA, vigorous PA

points for crude PA intensity categories. There are numerous different cut points available, all with different values although they propose to represent the same intensities from a physiological point of view.[2] The choice of cut points highly affects the outcome and make results difficult to compare between studies.[21] The detailed intensity spectrum used in our study is more transparent and does not depend on specific cut points. Because of the high detail, adjacent bins often represent the same physical behavior. Therefore, using different bin specifications would likely not influence the results.

The main advantage of PLS-DA compared to univariate statistical methods is that it provides measures of overall statistical uncertainty and strength of the pattern that discriminates the groups. The importance of specific variables (standardized and unstandardized) can be investigated with univariate statistics as well. Aadland et al. have previously shown that using simple bivariate correlation is able to show the association between a specific part of the PA intensity spectrum and cardio-metabolic health reasonably well without the use of multivariate statistics.[4] Similarly, simply displaying the group means and confidence intervals gives a good estimation of the pattern discriminating the groups. However, these methods do not provide accurate measures of variation and covariation and consequently do not describe uncertainty. Further, it is sounder to present the uncertainty of the entire model, instead of multiple *p*-values. Traditional group comparison requires one hypothesis test for each intensity variable, which would increase the risk of type I error. Another option for multivariate group comparison is multivariate analysis of variance (MANOVA). MANOVA is able to handle collinearity to some extent, but should not be used if the collinearity is larger than 90%.[22] Since the collinearity between the variables representing the PA intensity spectrum often is above 90%, this method is not suitable for these group comparisons. Other potential methods that are able to deal with the apparent collinearity is, for example ridge regression or lasso regression. Future studies should compare the performance of different methods of multivariate analysis.

Both the standardized and unstandardized pattern from the PLS-DA analysis suggest that the intervention group were less active in the MPA-VPA transition. However, the intervention group's higher volume of high end VPA was only shown in the standardized pattern. This is explained by this difference being small in absolute terms, but large in relation to the within groups variation at this intensity. Although the unstandardized pattern is similar between the multivariate and univariate analysis, the univariate pattern provides an absolute difference in minutes instead of multivariate coefficients and is therefore easier to interpret. The difference in the highest part of VPA consists of <1 min per day difference between groups for each variable. This difference might seem too small to be relevant. However, since the variables are highly collinear at this intensity, the pattern should be interpreted by aggregating the difference from 10 000 and above. This would subsequently represent a difference of a few minutes per day between groups, which could be considered more relevant. The multivariate analysis considers the entire PA intensity spectrum simultaneously, and therefore the difference in a single intensity variable cannot be isolated.[19] Furthermore, the small absolute group difference in the highest part of

VPA also explains why the traditional cut-point analysis suggests that there is a tendency toward the control group performing more VPA overall. The MPA to VPA transition is equivalent to brisk walking, whereas above 10 000 CPM is equivalent to high-intensity running.[23] According to the intervention protocol, the intervention group would have performed more PE classes consisting of ball games, running and jumping.[9] Therefore, the apparent difference PA intensity pattern identified in our study may reflect the intention of the intervention.

The group difference in LPA intensity is apparent in the unstandardized pattern from the PLS-DA analysis as well as in the standardized and unstandardized pattern from the univariate analysis, but not in the standardized pattern from PLS-DA. This discrepancy is likely a result of the target projection procedure involved in calculating the proportion of explained variance that represents the standardized pattern, since this is an additional procedure involved in standardizing the PLS-DA coefficients. This suggests that although there seems to be a difference in average LPA volume, this difference might be less influential in discriminating the groups. The explained variance of the pattern, also referred to as selectivity ratio, is standardized with regard to its predictive performance rather than its variation.[18] Therefore, the explained variance visualized in the figures should not be compared between PLS-DA models without taking the overall strength of the model into consideration. A specific variable on the PA intensity spectrum could be highly influential in relation to the other variables, but the overall discrimination strength of the model could still be small.

Physical activity data must be interpreted with regard to its compositional nature. Since all days are limited to 24 h, an increase in time spent at one intensity level must be accompanied by a decrease in time spent at another intensity level.[24] Therefore the discrimination patterns not only include at what PA intensities the intervention group spent more time, but also include from what intensities time was reduced. Consequently, the sum of the PLS-DA coefficients multiplied by the mean PA pattern of all study participants is zero. Similarly, the sum of the selectivity ratio multiplied by the mean standardized PA pattern this is also zero. The same applies to the univariate group differences. This implies that a single level of the intensity spectrum cannot be isolated but must be interpreted together with the entire discriminative PA pattern simultaneously.[19]

Partial least squares analysis is a well-established statistical method that is able to handle completely collinear variables.[6,25] In PA research it has been shown to correctly handle the multicollinearity apparent in variables representing PA intensity spectrum.[3] Although PLS was not originally developed for group comparison, PLS-DA has been shown to be a highly useful tool for discrimination purposes with multiple collinear variables both from a practical and theoretical point of view.[7] When understanding the basic utility of PLS-DA, the next step would be to expand to more complex models, investigating longitudinal change of the PA pattern in multiple groups by applying more components in multilevel designs (e.g. clustering, random effects etc.).

The predictive strength of the models in the current study was assessed by the proportion of correctly classified individuals. Most often, the proportion of explained variance to total variance (R^2) is used to describe the strength of PLS models or regression models in general. However, for R^2 to be a meaningful metric, the output must be continuous. In the current case, with a discrimination problem analysis, the performance should be assessed by a metric that takes the dichotomous nature of the response variable into consideration.[14] Although there are other metrics that could be used for this purpose, proportion of correctly classified individuals was chosen since it is easily understandable from a group comparison perspective.

There is no consensus on how to determine the optimal number of components.[15] Many different approaches have been suggested including Monte Carlo resampling, cross-validation and bootstrapping, sometimes nested in multiple loops.[14,15] The main part of the predictive performance usually comes from the first few components. This is explained by the PLS algorithm finding the components that best explains variation in the predictor variable successively. With increasing number of components, the additional benefit decreases. In the current study, Monte Carlo resampling was applied together with a backwards selection procedure.[15] This implies that the number of components were selected not to find the strongest model, but to ensure a substantially stronger model compared to one with fewer components, based on a probability threshold. If a PLS model representing PA is overfitted, the direction of the selectivity ratio, that comes from the original PLS coefficients, typically changes rapidly between positive and negative multiple times.[3] If this is apparent with a validated PLS model, the probability threshold can be lowered from 0.401 (0.25 standard deviations) to 0.308 (0.5 standard deviations).[15] This validation technique allows the researcher to balance the risk of over- and underfitting the model and is consistent with other studies applying PLS in PA research.[4,5]

In the repeated measures analyses, the improvement of the group discrimination by the multilevel extension compared to the independent model is expected since between subjects variation is excluded and should be relevant when comparing the PA pattern between M1 and M2.[8] Similar to the results of the current study, children have previously been shown to spend more time sedentary and less time physically active with increased age.[26] To investigate the outcome of the intervention specifically, a baseline measurement of PA would have been desirable. With a baseline measurement available, a paired analysis comparing the groups with regard to individual difference in PA patterns between the repeated measures could be performed. In this case, the individual PA spectrum difference would be used as predictor variables and group belonging as response variables when applying PLS-DA. However, since the aim of the current study was to compare statistical methods rather than investigating the outcome of the added PE, the longitudinal follow-up design was suitable for the study aim and school clustering of children would not influence the results.

Because of technical limitations at the time of data collection, only vertical acceleration was collected in the current study. Although vertical acceleration seems to be the most important axis to capture PA, combining the three axes to a vector magnitude is stronger associated to energy expenditure and cardio-metabolic health outcomes.[27,28] Furthermore, adding all three axes separately in the PLS model yields a stronger relationship to cardio-metabolic health outcomes than the vector magnitude alone.[28] Still, only considering time in intensity categories does not cover all aspects of PA. Since PLS analysis is developed for using a high number of predictor variables, even more variables than axes specific intensity and volume could be included. Such variables that could be of importance is for example frequency, bouts and activity type.[1] A further limitation of the present study is that the ActiGraph method used for PA measurement tends to overestimate high-intensity PA.[29] At high intensity, 90% of the time can be misclassified. This misclassification is caused by the narrow frequency filter involved in the raw data processing, which removes information in the accelerometer signal related to high-intensity PA and makes it difficult to distinguish this intensity.

5 | PERSPECTIVE

This study applies detailed analysis of PA intensity spectrum in group comparison for the first time. The analysis method enables investigation of PA intensity patterns with very high resolution in various study designs and represents a necessary methodological progression for more complete use of accelerometer data. The independent group analysis could be used in clinical and epidemiological studies to compare subgroups, or in intervention studies to compare baseline data of an intervention group to a control group. In addition, the multilevel extension could be applied in longitudinal studies, either intervention or epidemiological, to investigate change over time, as well as in studies with a crossover design.

ACKNOWLEDGEMENTS

This research was funded by Lund University, grant number 46118, Skåne University Hospital, grant number 98913 and Region Skåne, grant number 819961.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Jonatan Fridolfsson https://orcid.org/0000-0002-7003-4025
Daniel Arvidsson https://orcid.org/0000-0003-2365-187X

REFERENCES

1. Arvidsson D, Fridolfsson J, Börjesson M. Measurement of physical activity in clinical practice using accelerometers. *J Intern Med*. 2019;286(2):137-153. https://doi.org/10.1111/joim.12908
2. Migueles JH, Cadenas-Sanchez C, Ekelund U, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med*. 2017;47(9):1821-1845. https://doi.org/10.1007/s40279-017-0716-0
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):74. https://doi.org/10.1186/s12966-019-0836-z
4. 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(1):77. https://doi.org/10.1186/s12966-018-0707-z
5. Fridolfsson J, Börjesson M, Ekblom-Bak E, Ekblom Ö, Arvidsson D. Stronger association between high intensity physical activity and cardiometabolic health with improved assessment of the full intensity range using accelerometry. *Sensors*. 2020;20(4):1118. https://doi.org/10.3390/s20041118
6. Wold S, Ruhe A, Wold H, Dunn WJ III. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J Sci Stat Comput*. 1984;5(3):735-743. https://doi.org/10.1137/0905052
7. Barker M, Rayens W. Partial least squares for discrimination. *J Chemom*. 2003;17(3):166-173. https://doi.org/10.1002/cem.785
8. Westerhuis JA, van Velzen EJJ, Hoefsloot HCJ, Smilde AK. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics*. 2010;6(1):119-128. https://doi.org/10.1007/s11306-009-0185-z

9. Dencker M, Thorsson O, Karlsson MK, et al. Daily physical activity in Swedish children aged 8–11 years. *Scand J Med Sci Sports*. 2006;16(4):252-257. https://doi.org/10.1111/j.1600-0838.2005.00486.x
10. Linden C, Ahlborg HG, Besjakov J, Gardsell P, Karlsson MK. A school curriculum-based exercise program increases bone mineral accrual and bone size in prepubertal girls: two-year data from the Pediatric Osteoporosis Prevention (POP) study. *J Bone Miner Res*. 2006;21(6):829-835. https://doi.org/10.1359/jbmr.060304
11. Tryon WW, Williams R. Fully proportional actigraphy: a new instrument. *Behav Res Methods Instrum Comput*. 1996;28(3):392-403. https://doi.org/10.3758/BF03200519
12. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, Mcdowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181-188. https://doi.org/10.1249/mss.0b013e31815a51b3
13. 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):1557-1565. https://doi.org/10.1080/02640410802334196
14. Szymańska E, Saccenti E, Smilde AK, Westerhuis JA. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*. 2012;8(1):3-16. https://doi.org/10.1007/s11306-011-0330-3
15. 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 Chemom*. 2018;32(4):e2993. https://doi.org/10.1002/cem.2993
16. Lindgren F, Hansen B, Karcher W, Sjöström M, Eriksson L. Model validation by permutation tests: applications to variable selection. *J Chemom*. 1996;10(5–6):521-532. https://doi.org/10.1002/(SICI)1099-128X(199609)10:5/6<521:AID-CEM448>3.0.CO;2-J
17. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics*. 1994;138(3):963-971.
18. Kvalheim OM. Interpretation of partial least squares regression models by means of target projection and selectivity ratio plots. *J Chemom*. 2010;24(7–8):496-504. https://doi.org/10.1002/cem.1289
19. 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. https://doi.org/10.3390/metabo9070129
20. Alwis G, Linden C, Stenevi-Lundgren S, et al. A school-curriculum-based exercise intervention program for two years in pre-pubertal girls does not influence hip structure. *Dyn Med*. 2008;7(1):8. https://doi.org/10.1186/1476-5918-7-8
21. Fröberg A, Berg C, Larsson C, Boldemann C, Raustorp A. Combinations of epoch durations and cut-points to estimate sedentary time and physical activity among adolescents. *Meas Phys Educ Exerc Sci*. 2017;21(3):154-160. https://doi.org/10.1080/1091367X.2017.1309657
22. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 6th ed. Pearson New International Edition; 2014.
23. Trost SG, Ward DS, Moorehead SM, Watson PD, Riner W, Burke JR. Validity of the computer science and applications (CSA) activity monitor in children. *Med Sci Sports Exerc*. 1998;30(4):629-633. https://doi.org/10.1097/00005768-199804000-00023
24. 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):e0139984. https://doi.org/10.1371/journal.pone.0139984
25. Mehmood T, Liland KH, Snipen L, Sæbø S. A review of variable selection methods in partial least squares regression. *Chemom Intell Lab Syst*. 2012;118:62-69. https://doi.org/10.1016/j.chemolab.2012.07.010
26. Steene-Johannessen J, Hansen BH, Dalene KE, et al. Variations in accelerometry measured physical activity and sedentary time across Europe – harmonized analyses of 47,497 children and adolescents. *Int J Behav Nutr Phys Act*. 2020;17(1):38. https://doi.org/10.1186/s12966-020-00930-x
27. Santos-Lozano A, Santín-Medeiros F, Cardon G, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. *Int J Sports Med*. 2013;34(11):975-982.
28. 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):2173-2179. https://doi.org/10.1249/MSS.0000000000002021
29. Fridolfsson J, Börjesson M, Buck C, et al. Effects of frequency filtering on intensity and noise in accelerometer-based physical activity measurements. *Sensors*. 2019;19(9):2186. https://doi.org/10.3390/s19092186

How to cite this article: Fridolfsson J, Arvidsson D, Andersen LB, et al. Physical activity spectrum discriminant analysis—A method to compare detailed patterns between groups. *Scand J Med Sci Sports*. 2021;00:1–10. https://doi.org/ 10.1111/sms.14052