

BRIEF REPORT

Open Access



Analysing time-use composition as dependent variables in physical activity and sedentary behaviour research: different compositional data analysis approaches

Philip von Rosen^{1*}

Abstract

Recently, there has been a paradigm shift from considering physical activity and sedentary behaviour as “independent” risk factors of health to acknowledging their co-dependency and compositional nature. The focus is now on how these behaviours relate to each other rather than viewing them in isolation. Compositional data analysis (CoDA) is a methodology that has been developed specifically for compositional data and the number of publications using CoDA in physical activity and sedentary behaviour research has increased rapidly in the past years. Yet, only a small proportion of the published studies in physical activity and sedentary behaviour research have investigated the time-use composition as dependent variables. This could be related to challenges regarding the interpretation of the results and the lack of guidelines for deciding which statistical approach to use. Therefore, in this paper, four different approaches for analysing the time-use composition as dependent variables are presented and discussed. This paper advocates that the aim of research should guide how the dependent variable is defined and which data analysis approach is selected, and it encourages researchers to consider analysing time-use components as dependent variables in physical activity and sedentary behaviour research.

Keywords Compositional data analysis, Ilr coordinates, Isometric log-ratio, Outcome

Background

Traditionally, epidemiologists have explored time spent in different movement behaviours across a day, such as time spent in sleep, sedentary behaviour, light physical activity and moderate-to-vigorous physical activity, in relation to health outcomes. Recently, there has been a paradigm shift from considering movement behaviours as “independent” risk factors of health to acknowledging

their co-dependency and compositional nature [1–3]. Instead of exploring the health effects of these behaviours in isolation or with partial adjustments for one another, the importance of adequately adjusting for all parts of the time-use composition has been highlighted [2, 3]. In this perspective, time spent in movement behaviours can be seen as a composition, a vector of parts. A composition is a series of positive values (e.g. amounts of time spent in different movement behaviours), that adds up to the total available time in a given period (e.g. day, week, month) [4].

Given that the amounts of time spent in movement behaviours are parts of a finite total, these variables are perfectly multicollinear [4]. Therefore, a change in time

*Correspondence:

Philip von Rosen
philip.von.rosen@ki.se

¹Department of Neurobiology, Care Sciences, and Society (NVS) Division of Physiotherapy, Karolinska Institutet, Alfred Nobels Allé 23, Huddinge SE-141 83, Sweden



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

spent in one behaviour displaces time spent in, at least, one of the remaining behaviours in the composition. The compositional parts are commonly expressed as proportions (i.e. relative to the total time). This means they are bounded from zero to one (or 1 to 100 for percentages) and exist in a non-Euclidean space [4]. Compositional data analysis (CoDA) is a methodology that has been developed specifically for compositional data, and the number of publications using CoDA in physical activity and sedentary behaviour research has increased rapidly in the past several years. Analysing time-use compositions using statistical techniques that have been designed for non-compositional data may lead to misleading inferences [5, 6]. Therefore, several methodological papers have recommended to use CoDA when analysing time-use compositions in physical activity and sedentary behaviour research [2, 3, 6].

A widely used CoDA approach starts with a log-ratio transformation of compositional parts. While there are several types of log-ratios available, isometric log-ratios (*ilr*) [7] have been most used in physical activity and sedentary behaviour research. Applying the *ilr* transformation on a composition results in one less *ilr* coordinate compared to the number of compositional parts (e.g. two *ilr* coordinates will be calculated from a 3-part composition) [7]. The first *ilr* coordinate (ilr_1) represents time spent in one behaviour relative to the remaining behaviours. After this transformation, standard statistical techniques can be applied to the *ilr* coordinates [8].

In recent years, CoDA has drawn significant attention from researchers in the field of physical activity and sedentary behaviour research. A simple search through titles and abstracts of articles indexed in PubMed using the following syntax “compositional data analysis” AND “physical activity” shows a rapid increase in the number of publications in recent years. The first study in this field was published in 2015, followed by zero, four, 16 and 22 studies the following years to 2019. In 2020–2022, 37 to 41 studies have been published yearly. For example, CoDA has been applied to explore the association between movement behaviours and all cause-mortality, non-communicable diseases, biomarkers and movement skills [9–12]. However, surprisingly few studies have applied CoDA when analysing movement behaviours as dependent variables. Of the 151 studies published in 2015–2022, based on the above-mentioned literature search, only around 12% ($n=18$) analysed movement behaviours as the dependent variables [13–30]. This is a surprisingly small proportion of the publications, given that a common aim in epidemiological research is to explore changes in movement behaviours across time, differences in movement behaviours between groups or determinants for movement behaviours [31]. It has been much more common to apply CoDA for analysing

movement behaviours as independent variables in regression models. By entering all *ilr* coordinates as independent variables in a regression model, the model attempt to account for all portions of time spent in each behaviour that add up to a finite time. Thereby accounting for their combined effect on the dependent variable. However, analysing the composition of movement behaviours as the dependent variables may not be as straightforward and several different approaches could be considered. In this paper, we first describe how a time-use composition can be defined. Secondly, four different statistical approaches are presented and discussed based on regression models; (i) the first pivot coordinate (ilr_1), (ii) the first *pivot* coordinate adjusted for the remaining *pivot* coordinate(s), (iii) all *ilr* coordinates separately analysed, (iv) all *ilr* coordinates “stacked” into a single variable.

Defining a time-use composition

Typically, the time-use composition is defined as pivot coordinates or balance coordinates in physical activity and sedentary behaviour research. Using pivot coordinates, the coordinates are defined as one movement behaviour relative to the remaining behaviours. This approach has been used and described in several studies [3, 4, 10, 11]. Balance coordinates are a more flexible way of defining the time-use composition, focusing on the relationship between different parts of the movement behaviours. Sequential binary partition (SBP) can be used to partition movement behaviours into groups, leading to interpretable coordinates of the compositional data [32]. SBP defines an orthonormal basis that splits the composition into a series of non-overlapping groups (in our case groups of movement behaviours). This process is then repeated until each of the movement behaviours forms a single-part “group” (Table 1). Consequently, an SBP consists of groups with different combinations of movement behaviours and the coordinates can be seen as contrasts that are interpretable as one or several movement behaviours relative to one or several other behaviours. After the groups of parts are defined, which can be visualized using a balance dendrogram (Fig. 1), the *ilr* transformation can be applied. The new *ilr* coordinates (balances) can then be used as dependent variables. For example, a composition of time spent in sedentary behaviour, light, moderate and vigorous physical activity can be split into two groups: one including sedentary behaviour and light physical activity and the other including moderate and vigorous physical activity (ilr_1). The first group can then be split into time spent in sedentary behaviour relative to light physical activity (ilr_2) and the second group can be split into time spent in moderate physical activity relative to time spent in vigorous physical activity (ilr_3). Expert opinion or a data-driven approach to inform meaningful balances. Each of the balances can then be entered as a

Table 1 Example of sequential binary partition for four movement behaviours. The value “1” stands for inclusion in the nominator of a coordinate, the value “-1” stands for inclusion in the dominator of a coordinate, and the value “0” indicates that the part is not included to the coordinate

Coordinate	Coordinate 1 Sedentary behaviour & Light vs. Moderate & Vigorous	Coordinate 2 Sedentary behaviour vs. Light	Coordinate 3 Moderate vs. Vigorous
Sedentary behaviour	1	1	0
Light	1	-1	0
Moderate	-1	0	1
Vigorous	-1	0	-1

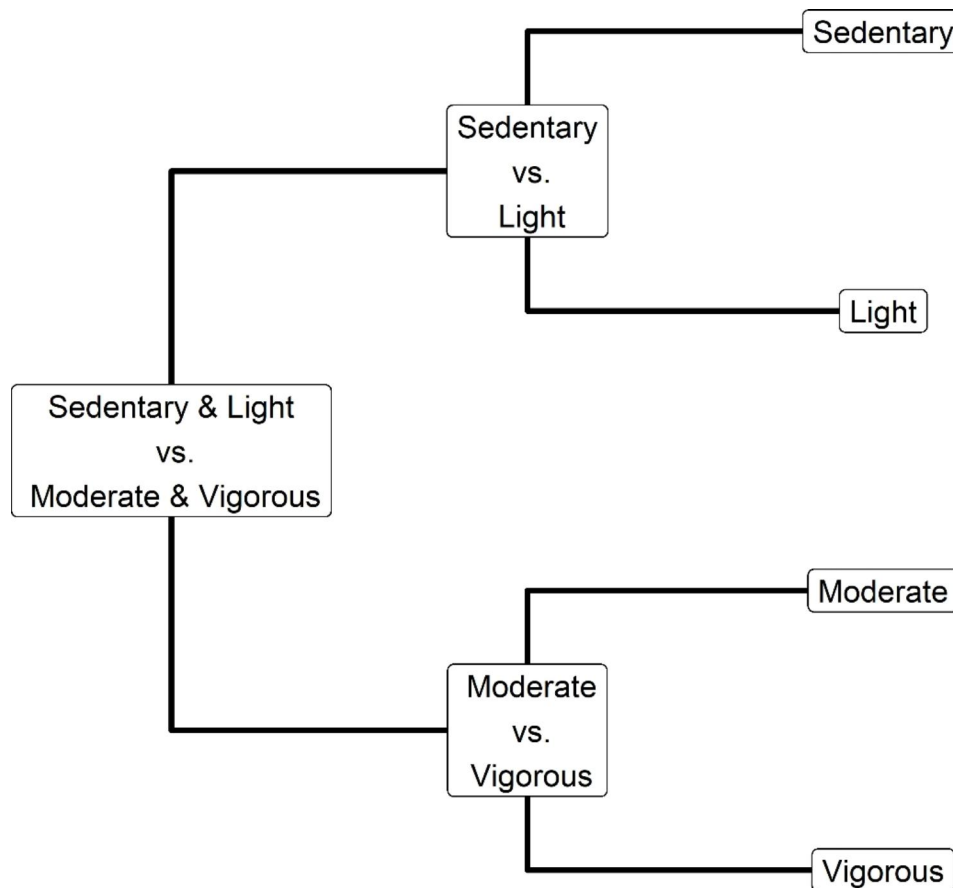


Fig. 1 Example of sequential binary partition of four movement behaviours, representing time spent sedentary behaviour, light, moderate and vigorous physical activity, contrasted in three coordinates. Note that the partitions results in non-overlapping groups

dependent variable in a regression model, with or without adjustments for the remaining balances. Balances has been used to explore time-use composition among preschoolers [12], changes in time-use composition during the transition to retirement [30] and followed a physical activity intervention [26]. The “Balance” package in R can be used for this purpose [33].

Approach I: the dependent variable is the first pivot coordinate (ilr_1)

Of the transformed coordinates, ilr_1 is the simplest and most interpretable variable to use as the dependent

variable. ilr_1 represents time in one behaviour relative to times in all the remaining behaviours. Since, the remaining ilr coordinates (i.e. $ilr_2, ilr_3, \dots, ilr_{D-1}$) do not contain the first part of the composition, we can interpret ilr_1 as the relative importance of one behaviour with respect to a (geometric) average of the remaining behaviours [3]. Consequently, if ilr_1 is used as the dependent variable, the model will simply explain which independent variables are associated with the dominance of one behaviour over the remaining behaviours. However, interpreting predicted values for ilr_1 may be difficult. By applying the inverse ilr transformation, time in a specific movement

behaviour can be obtained. Note that perturbations of compositional parts can be performed before the log-ratio transformation, so that the numerator in ilr_1 is time spent in any of the movement behaviours. For instance, if you are analysing a 3-part composition, you can create three pairs of ilr coordinates and analyse each ilr_1 separately. This approach has been used to evaluate the effect of a randomized trial on time-use composition among office workers [21], people with diabetes [34] and in a cross-sectional study explore the association between occupational and leisure time physical activity among workers [35].

Approach II: the dependent variable is the first pivot coordinate adjusted for the remaining pivot coordinate(s)

Before conducting a regression analysis, ilr_1 can be statistically adjusted for the remaining ilr coordinates, that is, the “influence” of the remaining ilr coordinates (i.e. $ilr_2, ilr_3, \dots, ilr_{D-1}$) can be statistically removed from ilr_1 . Thereby, it is possible to partially adjust or fully adjust ilr_1 . By regressing ilr_1 on ilr_2 (or other ilr coordinates), the unexplained variance of ilr_1 will be left, unexplained from ilr_2 . The residuals represent the variance of ilr_1 that is unexplained of the remaining ilr coordinates, and they can then be used as the dependent variable. This changes the interpretation of the model and may change the results of the analysis drastically, especially if ilr_2 is associated with ilr_1 in the final model. For instance, if age explains 60% of the variance of ilr_1 , adjusting ilr_1 for ilr_2 will likely reduce the explained variance if age is highly related to ilr_2 . If ilr_2 is not related to age, adjusting ilr_1 for ilr_2 might increase the percentage of explained variance. Instead of adjusting ilr_1 for the remaining ilr coordinates, it is also possible to include the ilr coordinates as independent variables in the regression model to account for their influence. Important to note is that it is likely more difficult to back-transform the data after statistical adjustment of one or several ilr coordinates. However, this approach might be useful when the study aims to explore in detail how independent variables are associated with a dependent variable, account for time spent in other movement behaviours.

Approach III: the dependent variables are separately analysed for all ilr coordinates

The ilr coordinates can be analysed in separate regression models, e.g. a regression model is performed for each ilr . By predicting values in each model, results in that predicted values for all ilr coordinates are estimated. The inverse ilr transformation can then be applied, resulting in that time in each movement behaviour being estimated. This approach has been used when exploring changes in time-use composition in a Thai population [22].

Approach IV: the dependent variables are all ilr coordinates “stacked” into a single variable

It is also possible to use all ilr coordinates as the dependent variable in a single regression model. This can be done by restructuring the dataset from a wide format to a long format, leading to all values for the ilr coordinates being placed in a single column. An indicator variable is then used to define which data belongs to which ilr coordinate. Therefore, there will be as many levels of the indicator variable as there are ilr coordinates. The indicator is then used as an independent variable in the model. By including interaction terms between the indicator variable and each of the independent variables in the model, the association with each ilr coordinate can be explored separately. However, as we have repeated measures for each participant (one value for each ilr) we need to use a model that can handle correlated observations, such as a mixed model that can contain both fixed and random effects. The advantage of this approach is that all ilr coordinates can be explored simultaneously in one model. As the dependent variable is based on several ilr coordinates, separate random effects for each ilr are possible to specify, which may increase the complexity of the model, while the accuracy of the model may improve. This approach has previously been used in a secondary analysis of a randomised trial, aiming to explore changes in medication load on 24-hour activity composition [36].

Methods

The publicly available NHANES 2003–2004 dataset was used to present examples of the four abovementioned approaches. Details on data collection in NHANES 2003–2004 have been described previously [37]. Participants were instructed to wear an ActiGraph 7164 accelerometer (ActiGraph, Shalimar, FL) on the right hip for seven consecutive days, to capture time in different intensities of physical activity. The device was set to sampling counts per 1-minute epochs and the *nhanesaccel* package for R (release 4.1.3; R Core Team, 2015, Vienna, Austria) was used to process the accelerometer data. In this analysis, data on time spent in sedentary behaviour (0–99 counts), light (100–759 counts), moderate (760–2019 counts) and vigorous physical activity (>2000 counts) were used as dependent variables and age of the participants was used as the independent variable. The age variable was centered by subtracting the mean (average) value, to facilitate interpretation of the coefficient. A valid day was defined as 10 or more hours of wearing an accelerometer and participants with records of 4 or more valid days were included in the data analysis. This resulted in a sample of 6340 participants.

Time spent in sedentary behaviour, light, moderate and vigorous physical activity was transformed into ilr coordinates. Given that a 4-part composition was used, each

movement behaviour was then represented by three *ilr* coordinates (z_1 , z_2 and z_3). Four different types of models were fitted where the independent variable was age and the dependent variable was (i) the first pivot coordinate (ilr_1), (ii) the first *pivot* coordinate adjusted for the remaining *pivot* coordinate(s), (iii) all *ilr* coordinates separately analysed, (iv) all *ilr* coordinates “stacked” into a single variable. To provide an example of balance coordinates, coordinates were defined as sedentary behaviour and light physical activity relative to moderate and vigorous physical activity (ilr_1), sedentary behaviour relative to light physical activity (ilr_2) and moderate physical activity relative to vigorous physical activity (ilr_3). For comparison purposes, a model with absolute time as the dependent variable was also fitted. In all models, linear regression was used, except for approach “iv”, where a linear mixed model was applied. The ilr_1 was back-transformed to the unbounded space by applying the inverse *ilr* transformation. As values for all *ilr* coordinates are needed in the *ilr* transformation, age-adjusted arithmetic mean values of ilr_2 and ilr_3 were calculated. In the model with absolute time as the dependent variable, the predicted values for all behaviours were transformed into percentages. R code for all analyses is provided as a [Supplementary File](#).

Results

In Table 2, the results for all models are shown. Age was significantly associated with absolute time and relative time in all movement behaviours in all models. Even if the point estimates differed across models, the overall results were similar, i.e. higher age was associated with lower time spent in light, moderate and vigorous physical activity, and higher time spent in sedentary behaviour. However, in approach III & IV higher age was associated with lower time spent in sedentary behaviour, and higher time spent in light physical activity.

The explained variance seemed to be higher in approach I & III compared with the model with absolute time as the dependent variable, for all dependent variables except moderate physical activity. The explained variance seemed to be lower in approach II, compared to approach I. Line plots based on fitted regression models of “absolute time”, approach I and approach III-V are presented in Figs. 2, 3, 4, 5 and 6. Note that the point estimates differed in Fig. 3A-D for different combinations of ilr_{1-3} , even if the trend of the regression lines was similar.

Discussion

Even if the overall results for the four different approaches were similar, our results illustrate that the results may vary depending on how the time-use composition is defined and treated in the analysis. In our case, the model with absolute time as the dependent variable

showed similar results as approach I-IV, demonstrating that age was associated with the absolute and relative time of different movement behaviours. Based on the coefficient of determination, approach I and III seem to explain the highest proportion of variance. However, the current study cannot state which model provides the most realistic estimates. Since the time-use composition is defined and treated differently across the approaches, direct comparisons are not possible. However, several considerations can be suggested and a summary of the approaches with comments is presented in Table 3.

How to define and treat the time-use composition should be based on the aim of the study. If a study aims at exploring the determinants of a specific movement behaviour relative to the remaining behaviours, approach I and II can be used. If the aim concerns determinants or between-group comparisons of time-use composition, approach III and IV can be used as estimates for all movement behaviours can be derived. Further, if a study explores determinants for a composition of multiple movement behaviours, the use of balance coordinates offers a systematic process to define the composition of movement behaviour. Obtaining predictive values can be used to interpret the models in all approaches, although it might be more difficult for approach II. Information criteria (e.g. Akaike’s information criterion, Bayesian information criterion, coefficient of determination) and parameter estimates can be used to guide which model to choose and how to interpret the model. Besides, the approaches described in this paper, several studies have used multivariate analysis of variance (MANOVA) to explore differences in time-use composition between groups [5, 13]. Regardless of which approach is chosen, the model assumptions still need to be checked despite how the time-use composition is defined and treated.

Conclusions

This paper advocates that the aim of research should guide how the dependent variable is defined and which data analysis approach is selected, and it encourages researchers to consider analysing time-use components as dependent variables in physical activity and sedentary behaviour research. Examples of different approaches for using the composition of time-use as dependent variables have been exemplified, discussed and codes for analyses have been provided. This is believed to strengthen the physical activity and sedentary behaviour research field and increase our understanding of movement behaviours co-dependency in various research settings.

Table 2 Regression models with age, centered, as independent variable and different combinations of time spent in sedentary behaviour, light, moderate and vigorous physical activity, as the dependent variable

Regression models				
Four linear regression models with absolute time as dependent variable				
Dependent variable	β	SE	P-value	R²
Sedentary behaviour	2.30	0.08	<0.001*	0.11
Light	-1.09	0.06	<0.001*	0.05
Moderate	-0.62	0.01	<0.001*	0.24
Vigorous	-0.05	0.002	<0.001*	0.08
Approach I: Four linear regression models with relative time (<i>ilr1</i>) as dependent variable				
Dependent variable	β	SE	P-value	R²
Sedentary behaviour_ <i>ilr1</i> ¹	0.04	0.001	<0.001*	0.38
Light_ <i>ilr1</i>	0.03	0.001	<0.001*	0.31
Moderate_ <i>ilr1</i>	-0.01	0.0003	<0.001*	0.13
Vigorous_ <i>ilr1</i>	-0.05	0.001	<0.001*	0.35
Approach II: Four linear regression models with relative time (unexplained variance of <i>ilr1</i>) as dependent variable				
Dependent variable	β	SE	P-value	R²
Sedentary behaviour_ <i>ilr1</i> ^a	0.01	0.0005	<0.001*	0.09
Light_ <i>ilr1</i> ^a	0.005	0.0003	<0.001*	0.04
Moderate_ <i>ilr1</i> ^b	-0.003	0.0003	<0.001*	0.02
Vigorous_ <i>ilr1</i> ^b	-0.009	0.0006	<0.001*	0.003
Approach III: Three linear regression models with all <i>ilr</i> coordinates analysed separately				
Dependent variable	β	SE	P-value	R²
Sedentary behaviour_ <i>ilr1</i>	0.04	0.0006001	<0.001*	0.38
Sedentary behaviour_ <i>ilr2</i>	0.04	0.0007001	<0.001*	0.35
Sedentary behaviour_ <i>ilr3</i>	0.03	0.0006001	<0.001*	0.26
Approach IV: One linear mixed model with values of <i>ilr1</i>, <i>ilr2</i> and <i>ilr3</i> as dependent variable				
Independent variable	β	SE	P-value	
Sedentary behaviour_ <i>ilr1</i>	3.28	0.01	<0.001*	
Light vs. moderate & vigorous_ <i>ilr2</i>	4.19	0.01	<0.001*	
Moderate vs. vigorous_ <i>ilr3</i>	3.11	0.01	<0.001*	
Age	0.04	0.01	<0.001*	
Interaction: Age* Light vs. moderate & vigorous_ <i>ilr2</i>	0.01	0.0004	<0.001*	
Interaction: Age* Moderate vs. vigorous_ <i>ilr3</i>	-0.01	0.0004	<0.001*	
Balance coordinates: Three linear regression models with <i>ilr</i> coordinates as dependent variable (created from sequency binary partition)				
Dependent variable	β	SE	P-value	R²
Sedentary behaviour & Light vs. Moderate & Vigorous_ <i>ilr1</i>	0.03	0.0005	<0.001*	0.35
Sedentary vs. Light_ <i>ilr2</i>	0.01	0.0002	<0.001*	0.11
Moderate vs. Vigorous_ <i>ilr3</i>	0.03	0.0006	<0.001*	0.26

^a Unexplained variance for time spent in moderate vs. vigorous physical activity
^b Unexplained variance for time spent in light vs. vigorous physical activity
ilr1, first isometric log-ratio coordinate; *ilr2*, second isometric log-ratio coordinate; *ilr3*, third isometric log-ratio coordinate; SE, standard error; β , slope, R², coefficient of determination.

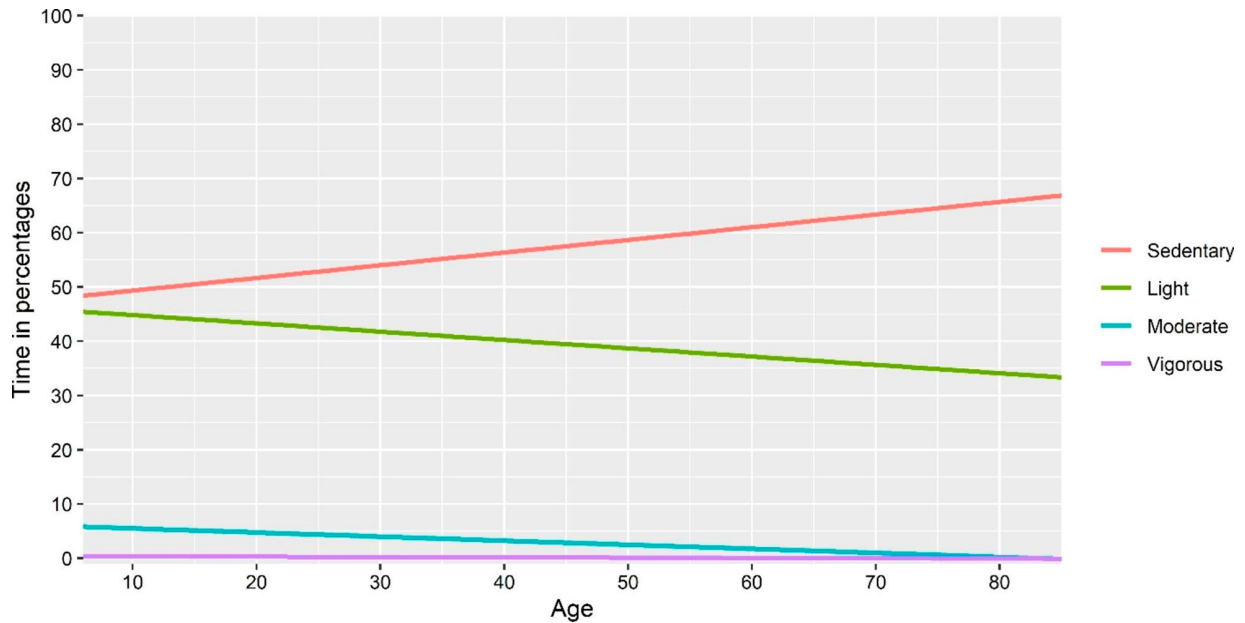


Fig. 2 Association between age and relative time in movement behaviours based on a linear regression model with absolute time in different behaviours as the dependent variable

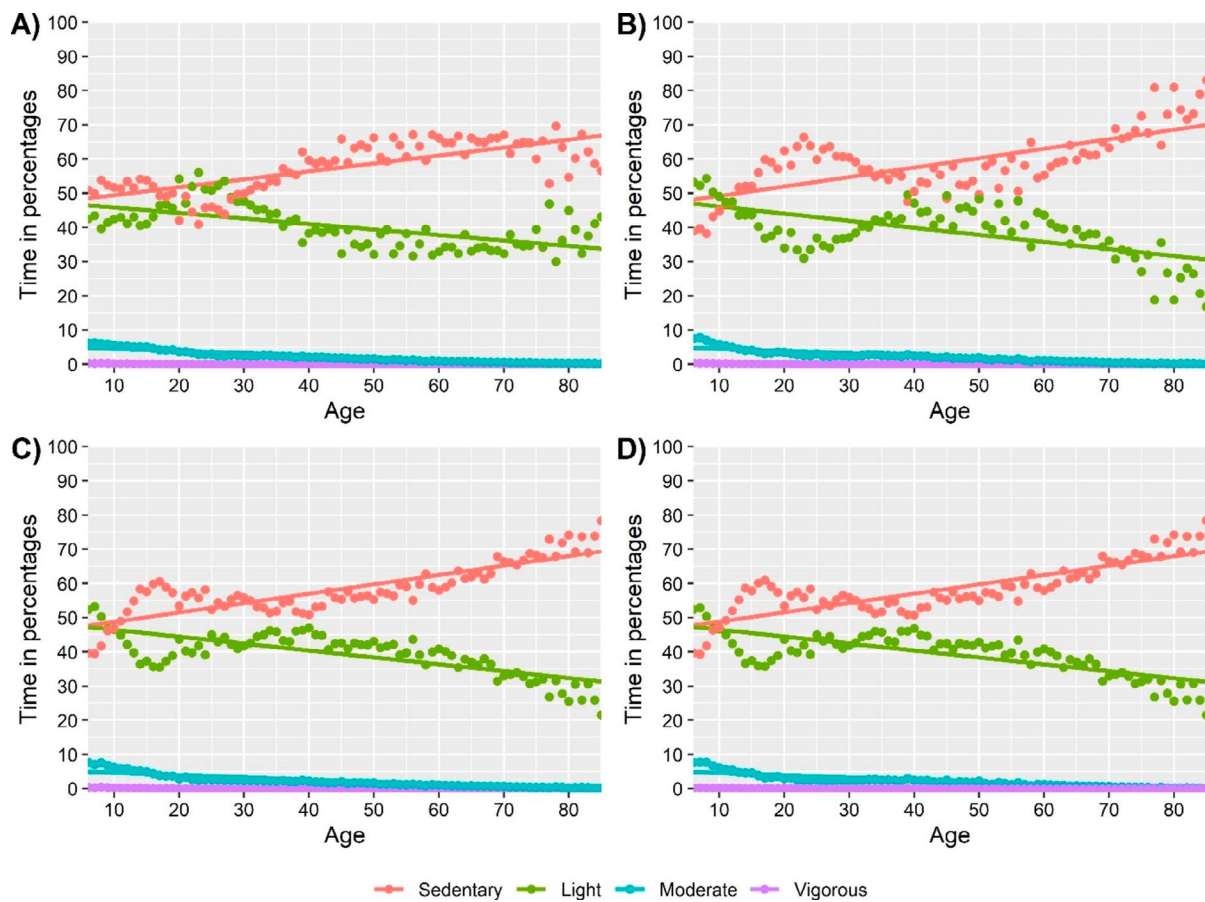


Fig. 3 Association between age and relative time in movement behaviours based on a linear regression model with the isometric log-ratio₁ coordinate for **A)** Sedentary behaviour, **B)** Light, **C)** Moderate, **D)** Vigorous physical activity as the dependent variable

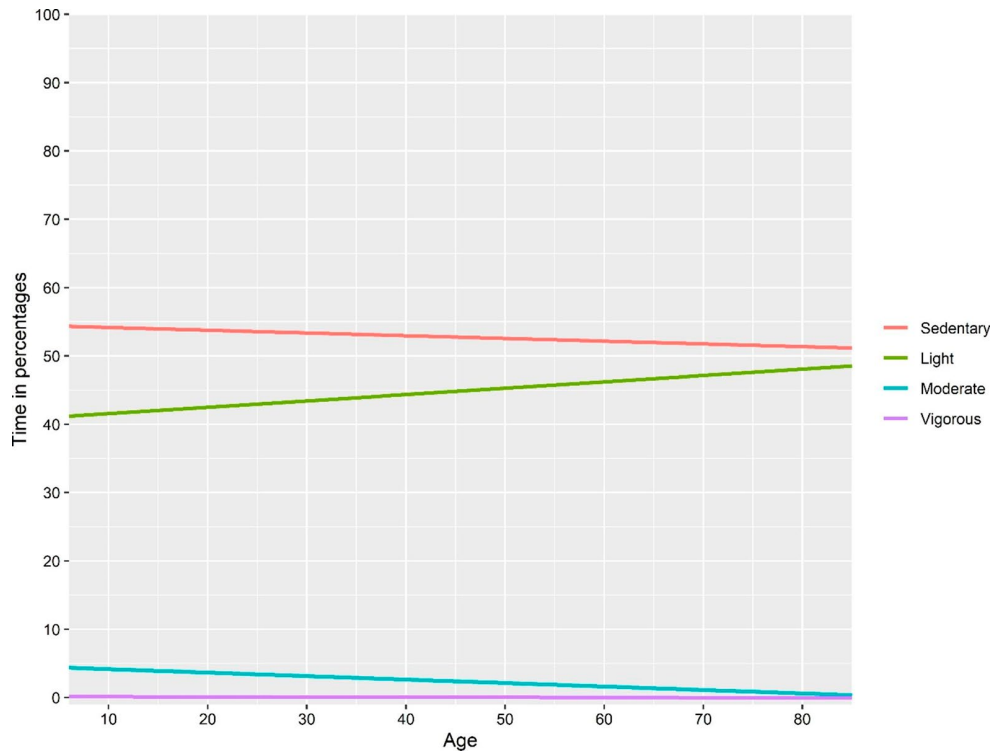


Fig. 4 Association between age and relative time in movement behaviours, where the isometric log-ratio₁₋₃ were separately analysed as the dependent variable in linear regression models

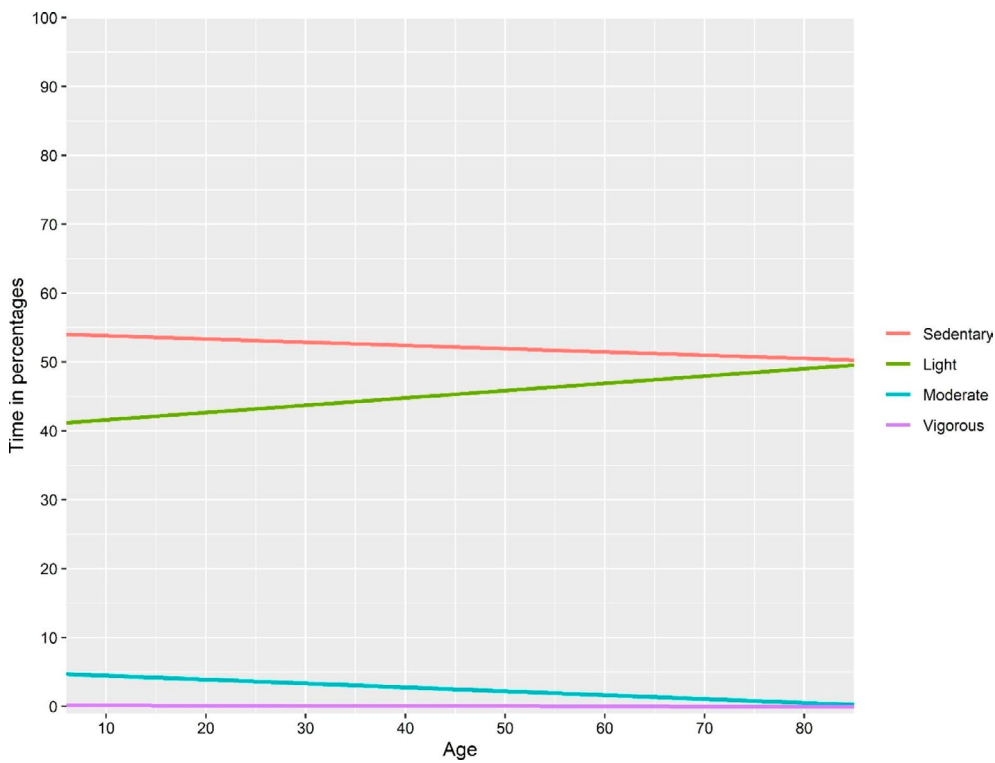


Fig. 5 Association between age and relative time in movement behaviours based on a linear mixed regression model, including the values of isometric log-ratio₁₋₃ coordinate as the dependent variable

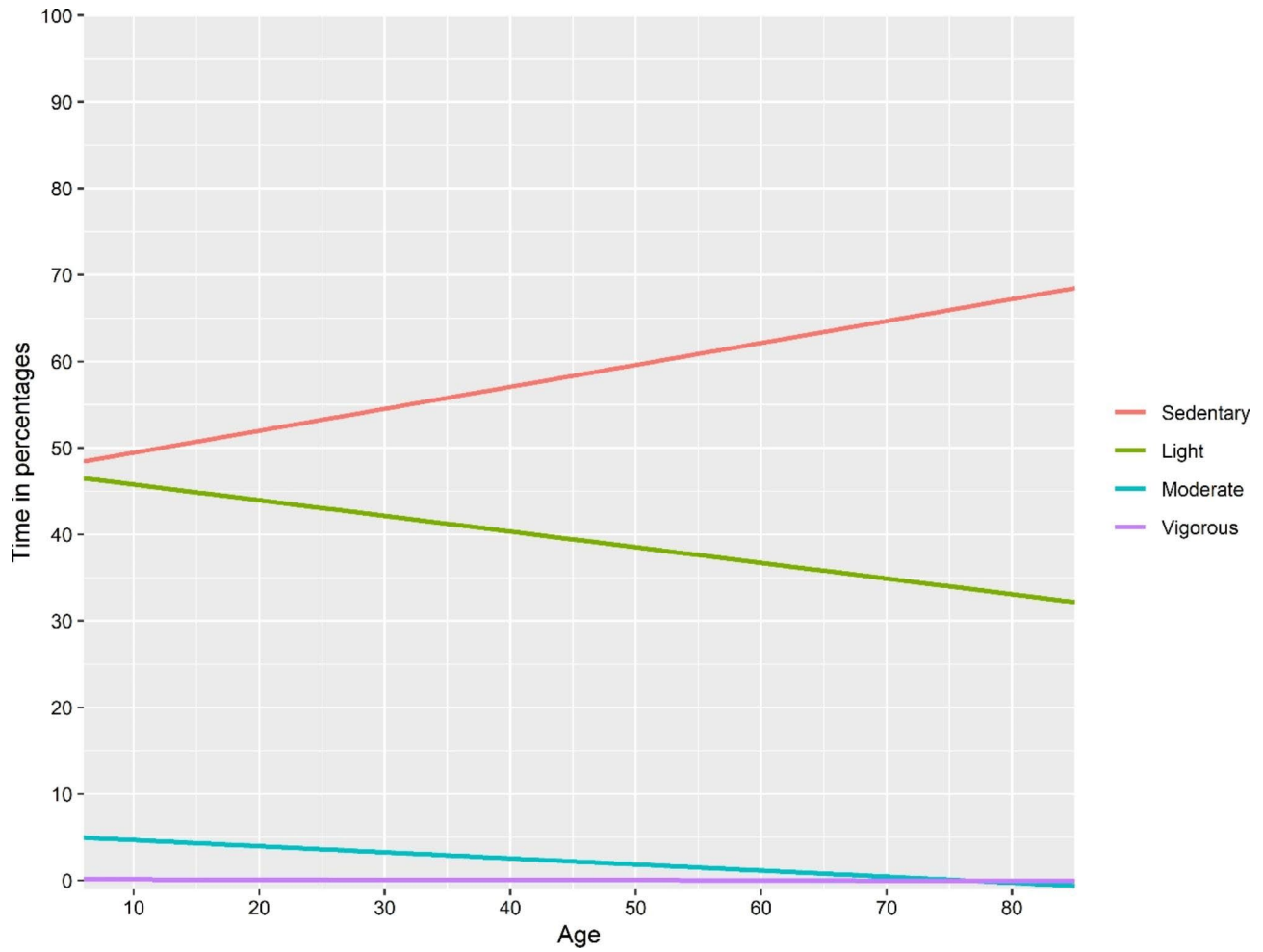


Fig. 6 Association between age and relative time in movement behaviours based on three linear regression models, where the dependent variables are balances. The balance coordinates were sedentary behaviour and light physical activity relative to moderate and vigorous physical activity (isometric log-ratio₁), sedentary behaviour relative to light physical activity (isometric log-ratio₂) and moderate physical activity relative to vigorous physical activity (isometric log-ratio₃)

Table 3 Four approaches to include time-use composition as the dependent variable in regression models with comments

Approach	Pros	Cons
Approach I: llr_1 is the dependent variable	Easy to interpret based on predicted values, information criteria and parameter estimates.	Ignore the remaining llr coordinates.
Approach II: The dependent variables is the unexplained variance (from other llr coordinates) of llr_1	Possible to statistically adjust for one or several llr coordinates. Information criteria and parameter estimates can be used to assess the model.	Difficult to interpret predicted values.
Approach III: All llr coordinates separately analysed as the dependent variable	Results can be interpreted based on predicted values, information criteria and parameter estimates.	Use all llr coordinates, but in separate models.
Approach IV: All llr coordinates "stacked" as the dependent variable	Results can be interpreted based on predicted values, information criteria and parameter estimates.	Extra parameters are needed to be estimated to distinguish between llr coordinates.

llr , isometric log-ratio.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s44167-023-00033-5>.

Supplementary Material 1

Acknowledgements

Not applicable.

Authors' contributions

PvR wrote the main manuscript text, prepared the tables and figures, and reviewed the manuscript.

Funding

Open access funding provided by Karolinska Institute.

Data Availability

Data are available in a public, open-access repository and can be accessed at the <https://www.cdc.gov/nchs/nhanes/index.htm>. R codes are available as a supplementary file.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Received: 13 June 2023 / Accepted: 1 September 2023

Published online: 02 November 2023

References

- Rosenberger ME, Fulton JE, Buman MP, Troiano RP, Grandner MA, Buchner DM, Haskell WL. The 24-Hour activity cycle: a new paradigm for physical activity. *Med Sci Sports Exerc.* 2019;51:454–64.
- Pedišić Ž. 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.
- Chastin SF, 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:e0139984.
- Aitchison J. The statistical analysis of compositional data. *J R Stat Soc Ser B.* 1982;44:139–60.
- Gupta N, Mathiassen SE, Mateu-Figueras G, Heiden M, Hallman DM, Jørgensen MB, et al. A comparison of standard and compositional data analysis in studies addressing group differences in sedentary behavior and physical activity. *Int J Behav Nutr Phys Act.* 2018;15:53.
- Dumuid D, Stanford TE, Martín-Fernández JA, Pedišić Ž, Maher CA, Lewis LK, Hron K, Katzmarzyk PT, Chaput JP, Fogelholm M, Hu G, Lambert EV, Maia J, Sarmiento OL, Standage M, Barreira TV, Broyles ST, Tudor-Locke C, Tremblay MS, Olds T. Compositional data analysis for physical activity, sedentary time and sleep research. *Stat Methods Med Res.* 2018;27:3726–38.
- Egozcue JJ, Pawłowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C. *Math Geol.* 2003;35:279–300. Isometric Logratio Transformations for Compositional Data Analysis.
- Pawłowsky-Glahn V, Egozcue JJ. Geometric approach to statistical analysis on the simplex. *Stoch Environ Res Risk Assess (SERRA).* 2001;15:384–98.
- Roscoe CMP, Duncan MJ, Clark CCT. The 24-h Movement Compositions in Weekday, Weekend Day or Four-Day Periods differentially associate with Fundamental Movement Skills. *Child (Basel).* 2021;8:828.
- von Rosen P, Dohrn IM, Hagströmer M. Association between physical activity and all-cause mortality: a 15-year follow-up using a compositional data analysis. *Scand J Med Sci Sports.* 2020;30:100–7.
- Rossen J, Von Rosen P, Johansson UB, Brismar K, Hagströmer M. Associations of physical activity and sedentary behavior with cardiometabolic biomarkers in prediabetes and type 2 diabetes: a compositional data analysis. *Phys Sportsmed.* 2020;48:222–8.
- de Souza AA, Clark CCT, Mota J, Tassitano RM, Duncan MJ, Martins CML. Movement behaviors in short versus adequate nocturnal sleepers: a compositional analysis of preschoolers. *Am J Hum Biol.* 2022;34:e23694.
- Brusaca LA, Januario LB, Mathiassen SE, Barbieri DF, Oliveira RV, Heiden M, Oliveira AB, Hallman DM. Sedentary behaviour, physical activity, and sleep among office workers during the COVID-19 pandemic: a comparison of Brazil and Sweden. *BMC Public Health.* 2022;22:2196.
- Campbell JE, Janssen I. The neighborhood physical environment and the 24-hour movement behavior composition among children. *Int J Environ Health Res.* 2023;33:575–87.
- Clifford Astbury C, Foley L, Penney TL, Adams J. How does Time Use Differ between individuals who do more versus less Foodwork? A compositional data analysis of Time Use in the United Kingdom Time Use Survey 2014–2015. *Nutrients.* 2020;12:2280.
- Foley L, Dumuid D, Atkin AJ, Olds T, Ogilvie D. Patterns of health behaviour associated with active travel: a compositional data analysis. *Int J Behav Nutr Phys Act.* 2018;21:15:26.
- Foley L, Dumuid D, Atkin AJ, Wijndaele K, Ogilvie D, Olds T. Cross-sectional and longitudinal associations between active commuting and patterns of movement behaviour during discretionary time: a compositional data analysis. *PLoS ONE.* 2019;16:e0216650.
- Hallman DM, Mathiassen SE, van der Beek AJ, Jackson JA, Coenen P. Calibration of self-reported Time spent sitting, standing and walking among Office Workers: a compositional data analysis. *Int J Environ Res Public Health.* 2019;27:16:3111.
- Hedayatrad L, Stewart T, Paine SJ, Marks E, Walker C, Duncan S. Sociodemographic differences in 24-hour time-use behaviours in New Zealand children. *Int J Behav Nutr Phys Act.* 2022;19:131.
- Johansson E, Mathiassen SE, Lund Rasmussen C, Hallman DM. Sitting, standing and moving during work and leisure among male and female office workers of different age: a compositional data analysis. *BMC Public Health.* 2020;20:826.
- Larisch LM, Bojsen-Møller E, Nooijen CFJ, Blom V, Ekblom M, Ekblom Ö, et al. Effects of two randomized and controlled multi-component interventions

- focusing on 24-Hour Movement Behavior among Office Workers: a compositional data analysis. *Int J Environ Res Public Health*. 2021;18:4191.
22. Liangruenrom N, Dumuid D, Pedisic Z. Physical activity, sedentary behaviour, and sleep in the Thai population: a compositional data analysis including 135,824 participants from two national time-use surveys. *PLoS ONE*. 2023;18:e0280957.
 23. Lund Rasmussen C, Palarea-Albaladejo J, Korshøj M, Gupta N, Nabe-Nielsen K, Holtermann A, Jørgensen MB. Is high aerobic workload at work associated with leisure time physical activity and sedentary behaviour among blue-collar workers? A compositional data analysis based on accelerometer data. *PLoS ONE*. 2019;614:e0217024.
 24. Olds T, Maher C, Dumuid D. Life on holidays: differences in activity composition between school and holiday periods in Australian children. *BMC Public Health*. 2019;19:450.
 25. Orme MW, Clague-Baker NJ, Richardson M, Drewry S, Robinson TG, Singh SJ. Does cardiac rehabilitation for people with stroke in the sub-acute phase of recovery lead to physical behaviour change? Results from compositional analysis of accelerometer-derived data. *Physiotherapy*. 2020;107:234–42.
 26. Pasanen J, Leskinen T, Suorsa K, Pulakka A, Virta J, Auranen K, Stenholm S. Effects of physical activity intervention on 24-h movement behaviors: a compositional data analysis. *Sci Rep*. 2022;12:8712.
 27. Pelclová J, Štefelová N, Olds T, Dumuid D, Hron K, Chastin S, et al. A study on prospective associations between adiposity and 7-year changes in movement behaviors among older women based on compositional data analysis. *BMC Geriatr*. 2021;21:203.
 28. Starbek P, Kastelic K, Šarabon N. The impact of online-schooling during COVID-19 on device-measured 24-Hour Movement Behaviours among High School students: a compositional data analysis. *Child (Basel)*. 2022;9:667.
 29. Štefelová N, Dygrýn J, Hron K, Gába A, Rubín L, Palarea-Albaladejo J. Robust compositional analysis of physical activity and sedentary Behaviour Data. *Int J Environ Res Public Health*. 2018;14:15:2248.
 30. Suorsa K, Leskinen T, Pasanen J, Pulakka A, Myllyntausta S, Pentti J, Chastin S, Vahtera J, Stenholm S. Changes in the 24-h movement behaviors during the transition to retirement: compositional data analysis. *Int J Behav Nutr Phys Act*. 2022;19:121.
 31. Pedišić Ž, Dumuid D, Olds S. Integrating sleep, sedentary behaviour, and physical activity research in the emerging field of time-use epidemiology: definitions, concepts, statistical methods, theoretical framework, and future directions. *Kinesiology: Int J Fundamental Appl Kinesiol*. 2017;49(2):252–69.
 32. Egozcue JJ, Pawłowsky-Glahn V. Groups of parts and their balances in compositional data analysis. *Math Geol*. 2005;37:795–828.
 33. Quinn T. Visualizing balances of compositional data: a new alternative to balance dendrograms. *F1000Research*. 2018;7:1278.
 34. Larsson K, Von Rosen P, Rossen J, Johansson U-B, Hagströmer M. Relative time in physical activity and sedentary behaviour across a 2-year pedometer-based intervention in people with prediabetes or type 2 diabetes: a secondary analysis of a randomised controlled trial. *JASSB*. 2023;2:10.
 35. Rasmussen CL, Palarea-Albaladejo J, Bauman A, Gupta N, Nabe-Nielsen K, Jørgensen MB, et al. Does physically demanding work hinder a physically active lifestyle in low socioeconomic workers? A compositional data analysis based on Accelerometer Data. *Int J Environ Res Public Health*. 2018;15:1306.
 36. Lim R, Dumuid D, Parfitt G, Stanford T, Post D, Bilton R, et al. Using wrist-worn accelerometers to identify the impact of medicines with anticholinergic or sedative properties on sedentary time: a 12-month prospective analysis. *Maturitas*. 2023;172:9–14.
 37. Johnson CL, Dohrmann SM, Burt VL, Mohadjer LK. National health and nutrition examination survey: sample design, 2011–2014. *Vital Health Stat 2*. 2014;162:1–33.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.