



HAL
open science

Predicting recovery after stressors using step count data derived from activity monitors

Dario Baretta, Sarah Koch, Joren Buekers, Judith M. Garcia-Aymerich, Lenka Knapova, Steriani Elavsky, Job Godino, Merlijn Olthof, Anna Lichtwarck-Aschoff, Ruud den Hartigh, et al.

► **To cite this version:**

Dario Baretta, Sarah Koch, Joren Buekers, Judith M. Garcia-Aymerich, Lenka Knapova, et al.. Predicting recovery after stressors using step count data derived from activity monitors. *npj Digital Medicine*, 2025, 8 (1), pp.606. <10.1038/s41746-025-01998-0>. <hal-05322598>

HAL Id: hal-05322598

<https://hal.science/hal-05322598v1>

Submitted on 20 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

<https://doi.org/10.1038/s41746-025-01998-0>

Predicting recovery after stressors using step count data derived from activity monitors

Check for updates

Dario Baretta¹ ✉, Sarah Koch^{2,3}, Joren Buekers^{3,4}, Judith Garcia-Aymerich^{3,5}, Lenka Knapova⁶, Steriani Elavsky⁶, Job Godino^{7,8}, Merlijn Olthof⁹, Anna Lichtwarck-Aschoff⁹, Ruud den Hartigh⁹ & Guillaume Chevance^{3,10}

This study examines the stressor-response process in physical activity among 226 participants across four countries. We analyzed their step count collected via activity monitors before and after a significant stressor: the COVID-19 lockdown. Results showed that a ‘local dynamic complexity’ metric significantly predicts the rate of recovery to pre-COVID levels of physical activity. These findings provide new opportunities for just-in-time interventions to support physical activity recovery after disruptive stressors.

Physical activity is key to health and wellbeing¹. However, this behavior is sensitive to numerous stressors which can perturb regular engagement. By stressors, we refer to external or internal events that can disrupt a system’s equilibrium, necessitating adaptive responses to restore balance^{2,3}. According to ecological models of physical activity⁴, stressors can occur at the intra-individual or extra-individual level. Intra-individual stressors originate from biological, psychological, or behavioral factors (e.g., illness or psychological distress)—whereas extra-individual stressors stem from the external environment, including the social, natural, and built physical settings, as well as policy regulations, all of which may hinder physical activity. Previous research—drawing from both broader resilience frameworks^{2,5,6} and empirical studies³—has suggested that individuals’ may respond differently to such perturbations: some may not return to prior physical activity levels (i.e., impairment), others may successfully recover (i.e., resilience), and some may even increase beyond their previous functioning (i.e., thriving, growth) (Fig. 1). Recent advances in mobile sensing have enabled precise characterization and quantification of these stressor-response trajectories through the analysis of high-resolution time series data^{2,5}. Among various outcomes, data from activity monitors, particularly step count, have emerged as powerful indicators of an individual’s physical state throughout the stressor-response trajectory (e.g.,⁷).

Step count is a passive digital biomarker that provides a scalable, unobtrusive, time-sensitive, and cost-effective way to continuously monitor individuals’ physical activity. While it does not capture every dimension or type of physical activity, it represents a direct and pragmatic indicator of daily movement patterns. Specifically, it enables researchers to capture key

features of the stressor-response trajectory, such as its immediate direction (e.g., increasing vs. decreasing) and momentary pace (Fig. 2). Direction and pace may vary over time, with reversals or temporary slowdowns in trajectory resulting from interactions with additional stressors that can operate at different levels. Tracking day-to-day changes in step count offers the opportunity to model and quantify the dynamics that govern how individuals respond to stressors. Ultimately, predicting future momentary direction and pace of the stressor-response trajectory (Fig. 2b–d) would enable the development and optimization of just-in-time intervention strategies—timely, targeted support provided at critical moments⁸—that could accelerate recovery and prevent delays or relapses that might occur in the weeks following the stressor.

While ecological models highlight the multilevel influences on physical activity, they offer limited tools for capturing the dynamic and nonlinear patterns of individual responses over time. Complex systems theory complements this perspective by providing a framework to model and predict individual progression throughout their response to stressors^{9–11}, particularly through early warning signals. Statistical markers of early warning signals include increases in autocorrelation, variance, and dynamic complexity (DC)^{11,12}. In particular, increased DC indicates the system’s heightened sensitivity to external influences (see “Methods” section) and has recently been demonstrated to be a predictor of future losses in step count^{13,14}.

Leveraging mobile sensing and grounded in the complex systems framework, this study aims to shed light on the largely unexplored stressor-response process. To achieve this, we conducted a multi-country study

¹Institute of Psychology, University of Bern, Bern, Switzerland. ²University of Basel, Basel, Switzerland. ³ISGlobal, Barcelona, Spain. ⁴CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. ⁵Universitat Pompeu Fabra (UPF), Barcelona, Spain. ⁶University of Ostrava, Ostrava, Czechia. ⁷Family Health Centers of San Diego, San Diego, CA, USA. ⁸UC San Diego, San Diego, CA, USA. ⁹Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands. ¹⁰Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail), Rennes, France. ✉e-mail: dario.baretta@unibe.ch

focusing on this process by analyzing the step counts of individuals following a significant stressor: the COVID-19 lockdown and its associated restrictions (e.g., enforced home confinement)¹⁵. Specifically, we investigated whether DC in step count, as measured by activity monitors, could predict the slowdowns (Fig. 2c vs. Fig. 2b) or even reversals (Fig. 2d vs. Fig. 2b–c) in the stressor-response trajectory, also derived from step counts. We hypothesized that increases in local DC would be associated with a decrease in the local rate of change dy in daily step counts in the following days.

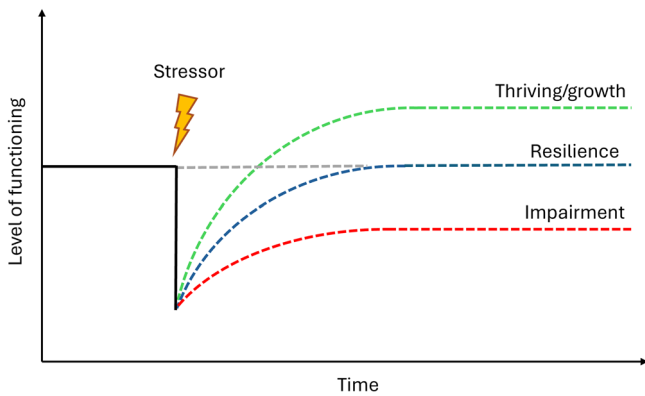


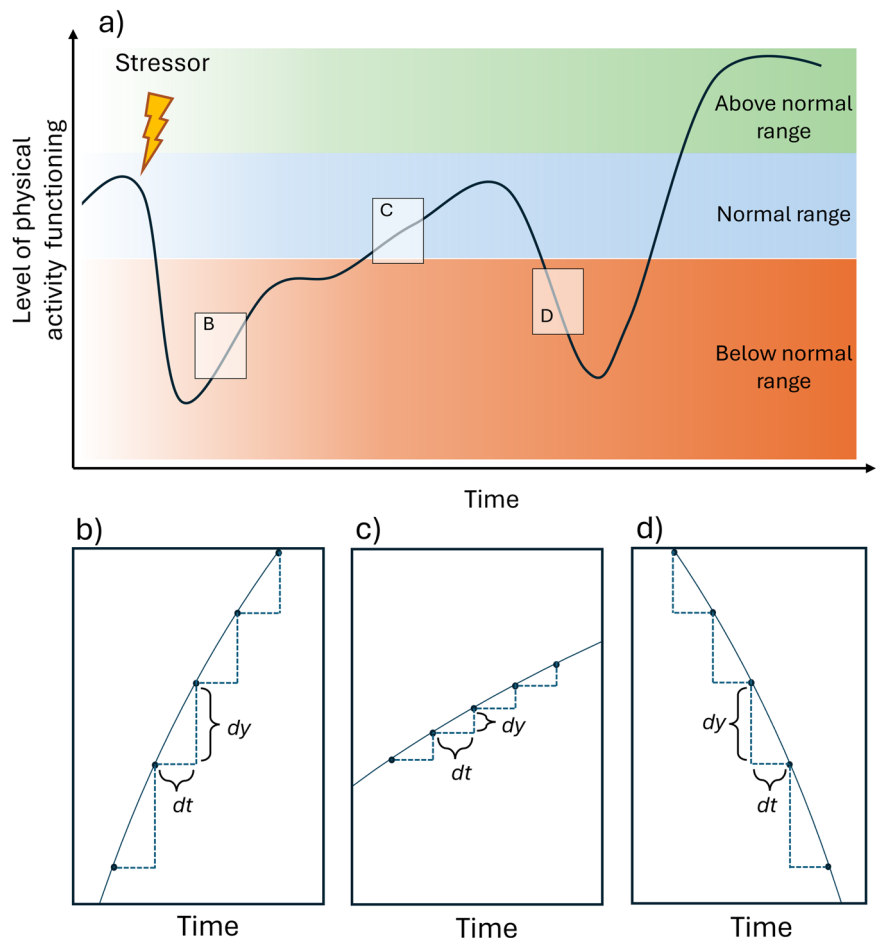
Fig. 1 | Potential responses to perturbations generated by a stressor. The black solid line represents a person’s state of functioning before encountering a stressor. Figure adapted from ref. 6. Copyright 2010 by John Wiley and Sons. Used with permission.

To test this hypothesis, we analyzed the stressor-response trajectory in daily step count following the first COVID-19 lockdown across 226 participants (77 from the United States, 65 from the Czech Republic, 48 from Spain, and 36 from Norway), resulting in a total of 44,825 daily step count observations from activity monitors (e.g., Fitbit, Garmin). Before the lockdown, the median daily step count was 10,167 (IQR: 8063–12,188; mean = 10,544), suggesting that the current sample was generally more active than the broader adult population and fell just below the highest quartile of daily step count reported in previous meta-analytic research¹⁶. During the 30 days following the lockdown, the median perturbation in daily steps was 3744 (IQR: 2312–5252; mean = 4063). Consequently, on average, participants shifted from taking over 10,000 steps per day to approximately 6500 steps per day, which falls below the recommended target for health benefits (e.g., 9000 steps per day,^{16,17}). For 110 participants (49%), it took a median of 88 days (IQR: 50–130; mean = 93) to return to pre-lockdown levels of daily steps, while 116 participants (51%) had not returned by September 30, 2020 (ca. 6.5 months after the first lockdown was issued). The average median rate of change in daily steps following the stressor was 11 steps per day, indicating that, in median, participants regained 11 steps per day as part of their response to the stressor (see Supplementary Table 1 for descriptive statistics of key variables by data source).

After computing individuals’ daily rate of change in step count, we computed a local DC score to characterize critical fluctuations in the step count time series¹⁸. Local DC (independent variable) was used to predict the rate of change in the stressor-response trajectory (dependent variable). Then, we explored the cross-correlation between local DC and the rate of change in daily step count to identify the lag with the highest correlation. The average cross-correlation was checked from a 1-day lag to a 30-day lag.

Fig. 2 | A physical activity stressor-response process characterized by varying direction and pace.

a Conceptual visualization of various responses to stressors, with the overall response trajectory characterized by varying momentary pace and direction. **b–d** The within-person progression throughout the response process can be characterized by different rates of change, indicating whether the momentary pace is fast or slow (**b** vs. **c**), positive or negative (**b**, **c** vs. **d**). The rate of change, dy , reflects how much the outcome variable (here, *step count*) changes in response to a one-unit change, dt , in the independent variable (here, *time*). The rate of change can be operationalized as the first-order derivative of a growth function fitting the outcome variable. It can take positive, negative, or zero values depending on whether the change is positive, negative, or null.



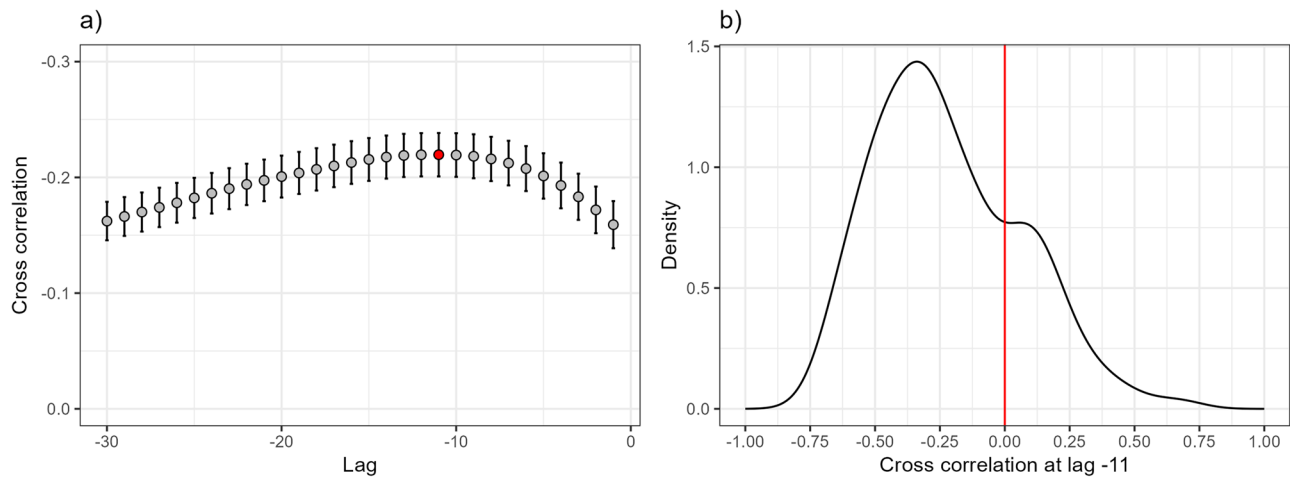


Fig. 3 | Cross-correlation between local DC and rate of change. **a** Shows the cross-correlation values between local DC and rate of change at various lags. **b** Shows the sample distribution of the cross-correlation values at lag-11.

Table 1 | Multilevel model summary

DV is rate of change	Fixed effects		Random effects
	Est. (SE)	95% CI	SD
Intercept	48.45 (4.89)	[38.85, 58.04]	70.89
Time	-43.51 (5.35)	[-53.98, -33.03]	78.42
Local DC	-20.42 (2.73)	[-25.78, -15.07]	39.85
Local DC ²	4.67 (1.14)	[2.44, 6.90]	16.11
Pre-lockdown median steps	-0.81(1.19)	[-3.15, 1.52]	-

Number observations = 39,401; ICC = 0.3; R2 Marg. = 0.067; R2 Cond. = 0.325

Values presented in bold are significant based on their 95% confidence intervals; R^2 Marginal = the part of the variance in the outcome explained by the fixed effects; R^2 Conditional = the part of the variance in the outcome explained by the full model.

Results, shown in Fig. 3, indicate that the highest average correlation occurred at lag-11 ($r = -0.22$).

After identifying the time lag at which local DC had its largest association with the rate of change, we conducted multi-level analyses to quantify the association between local DC at lag-11 and future rate of change in step count. The multilevel model included a random intercept for each participant and random slopes for three continuous independent variables: time, local DC, and the quadratic term of local DC. We found a negative effect of time ($B = -43.51$, 95% CI $[-53.98, -33.03]$), indicating that the rate of change decreased as more time passed since the stressor occurred (see Table 1). Local DC was a significant predictor of the rate of change, with a negative linear effect ($B = -20.42$, 95% CI $[-25.78, -15.07]$) and a positive quadratic effect ($B = 4.67$, 95% CI $[2.44, 6.90]$), suggesting that the negative association weakens at higher levels of local DC. Based on these estimates, the overall effect of local DC is -15.75 at one standard deviation above the mean, and 25.09 at one standard deviation below the mean (see Fig. 4 and Supplementary Fig. 1). To help contextualize the clinical relevance of these findings, it is useful to consider the distribution of local DC values across the observation window. Participants had, on average, 28 days with local DC values exceeding one standard deviation above the mean, which together predicted a cumulative delay of about 440 daily steps (28 days \times -15.75 steps) in the recovery trajectory by the end of the study period. Conversely, they had an average of 24 days with local DC values falling one standard deviation below the mean, which together predicted a cumulative acceleration of about 600 daily steps (24 days \times 25.09 steps).

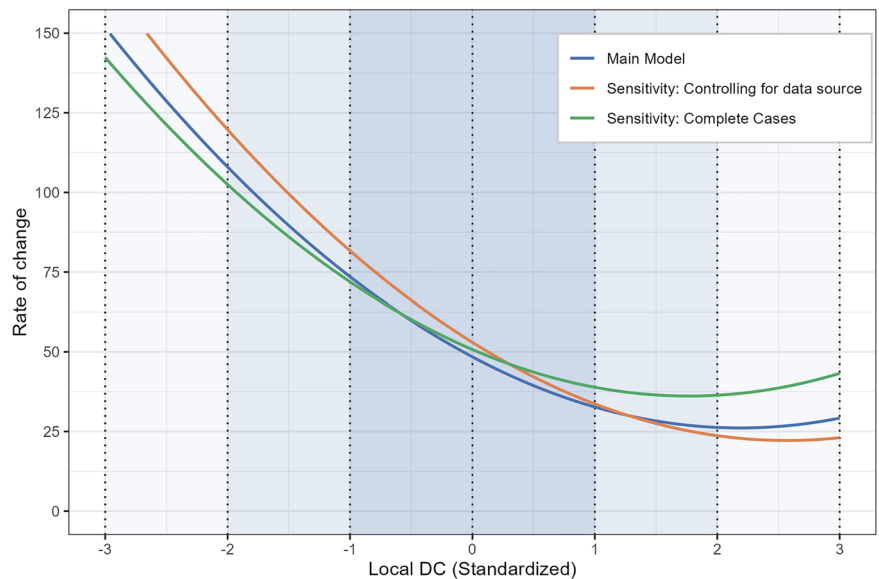
Sensitivity analyses were conducted to evaluate the robustness of the current findings (Fig. 4 and Supplementary Tables 2–6). We tested whether the association between local DC on the rate of change was moderated by the

data source (see Table 2 for an overview of included datasets). The results indicated that this association was stronger when controlling for the moderating effect of the United States-based Smart 2.0 study (i.e., the association was weaker in this data source). Specifically, the linear effect was $B = -24.01$ (95% CI $[-29.97, -18.05]$) and the quadratic effect was $B = 4.68$ (95% CI $[2.45, 6.91]$), resulting in a combined effect of -19.33 at one standard deviation above the mean and 28.69 at one standard deviation below the mean. An additional sensitivity analysis based solely on complete cases (i.e., excluding missing values) showed a reduced but still significant effect (combined estimate = -13.03 at one standard deviation above the mean and 20.00 at one standard deviation below the mean).

In summary, this study examines how step count data from activity monitors can characterize and predict the stressor-response trajectory in physical activity. Building on previous work^{13,14}, we found that local DC significantly predicted the rate of change in daily step count after a significant stressor (i.e., the COVID-19 lockdown). This association was non-linear, with lower local DC values being a stronger predictor of accelerated recovery than higher local DC values were of slowdowns (see Fig. 4). These effects are clinically meaningful, particularly when compared to the 3500 steps needed, on average, to return to pre-stressor levels in our sample or the roughly 2500 steps required to meet the recommended target of 9000 daily steps^{16,17}. Both elevated and reduced local DC may therefore function as early warning signals, respectively indicating increased risk of slowdowns or greater potential for recovery in the stressor-response trajectory in the following days.

When interpreting these findings, it is important to acknowledge some limitations. First, while aggregating data from multiple studies is a key strength of this work, it also introduces potential biases due to heterogeneity

Fig. 4 | Non-linear association between local DC and the rate of change in daily step count. The three lines represent the modeled non-linear association (combining linear and quadratic terms) between local DC and the rate of change in daily step count. The blue line corresponds to the main model (Table 1), the orange line to the model controlling for the moderating effect of the United States-based Smart 2.0 study (Supplementary Table 2), and the green line to the complete case analysis (Supplementary Table 4). The rate of change values at local DC = 0 reflect each model's specific intercept. The associations between local DC and rate of change are interpreted with the time effect held constant at zero.



in study-specific characteristics, such as study design, data collection protocols, and country-specific lockdown measures. Although we conducted a moderation analysis to address this issue, fully disentangling the factors underlying differences across studies (e.g., geographic location versus study design) was not possible. Moreover, harmonizing demographic characteristics that may have influenced the rate of change—such as socioeconomic status—was not feasible across studies. Second, the lack of wear time data in some datasets limited our ability to accurately identify and exclude invalid step count observations. Additionally, although the activity trackers used to assess step count (e.g., Fitbit, Garmin) have the advantage of being widely accessible and generally provide accurate step count metrics, they do not match the precision of gold-standard research tools^{19–21}, and their proprietary algorithms limit transparency and replicability. Finally, the relatively active nature of the sample may constrain the generalizability of the findings to less active populations. This characteristic likely stems from study-specific inclusion criteria (e.g., a portion of the sample recruited from a runner population)²², as well as from the fact that other participants were already using activity trackers^{23,24}, which is often associated with a greater propensity to pursue specific goals, such as improving health²⁵.

Altogether, the findings of this study underscore the potential of local DC to predict progression throughout the physical activity response to stressors. This metric fits particularly well in contexts where the individual response to major stressors (e.g., the lockdown) is shaped by interactions with additional influences at different levels (e.g., the work or family environment) or with intra-individual factors (e.g., reduced psychological resources), as suggested by ecological models of physical activity⁴. Importantly, local DC is derived entirely from sensor-based data and does not require additional participant resources, such as frequent questionnaires, which are often a barrier in long-term interventions and health surveillance. This early warning signal thus offers valuable opportunities for developing just-in-time interventions that provide timely support during critical stages in the recovery process⁸. For example, when local DC shows an increasing trend indicating instability, an intervention could be triggered to support planning of structured activity routines or initiate goal-setting strategies. However, future empirical work is warranted to evaluate whether just-in-time intervention strategies triggered by changes in local DC can improve recovery after stressors.

Replicating this approach in additional contexts where a successful response to stressors is crucial (e.g., after surgery, during cardiac rehabilitation, following chemotherapy) is necessary to validate these findings across diverse settings. Moreover, extending this approach to other physical activity signals (e.g., minutes of moderate-to-vigorous physical

activity) or integrating additional passive data—such as sleep, heart rate variability, GPS, or weather—with local DC derived from step count could help develop more specific early warning signals and enhance our ability to understand and predict stressor-response trajectories. For example, a concurrent rise in DC and physiological stress indicators (e.g., reduced heart rate variability) or disrupted sleep might signal a different type or severity of vulnerability than DC changes observed in isolation. Alternatively, these additional data streams could be combined to derive a multidimensional DC score, rather than calculating it solely based on step count. The increasing availability of such multimodal data presents a promising opportunity to develop more nuanced models of stressor response that integrate behavioral, physiological, and environmental dynamics—ultimately informing more tailored and effective intervention strategies.

Methods

Data sources

This study combined physical activity time series collected in the period between October 1st, 2019 and September 30th, 2020 as part of various studies as summarized in Table 2. Each study received ethical approval and written, informed consent to participate was obtained from all participants^{22–24,26}.

Location-specific COVID-19 context

Smart2.0 (San Diego, California, US). In response to COVID-19, a mandatory statewide stay-at-home order was implemented on March 19, 2020, resulting in the closure of gyms, beaches, parks, and recreational spaces. However, beaches and parks were reopened on April 27, 2020²⁷. In early May 2020, California launched a reopening program, with implementation delegated to individual counties. In late August, the state introduced a risk-based tier system to guide the tightening or easing of COVID-19 prevention measures in each county.

4HAIE (Czech Republic). The lockdown in Czechia began on March 16, 2020, with the implementation of restrictions on non-essential travel and the closure of borders with neighboring countries. Travel for work, family, healthcare, and necessities was permitted, as was outdoor physical activity, provided social distancing was observed. Restaurants, cafes, gyms, and most shops were closed. The lockdown officially ended on May 5, 2020, with restrictions being gradually lifted throughout the summer. However, starting in September, some restrictions, such as social distancing and masking, were reinstated incrementally²⁸.

Table 2 | Data sources information

Study		SMART 2.0 ²⁸	4HAIE ²²	COVICAT ²⁴	Open dataset ²³
Study characteristics	A 24-month (96 weeks), three-arm, parallel-group randomized controlled trial targeting weight loss among young adults with overweight or obesity.	A 12-month prospective cohort study, examining physical activity in an air-polluted environment.	A cohort study to characterize the health impact of the COVID-19 pandemic on the population. These data were retrospectively collected using crowdsourcing from participants of the cohort GCAT cohort ⁴⁵ .	An open dataset providing physical activity data during the COVID-19 pandemic. Two years of data were automatically downloaded from consenting participants by directly accessing the Garmin and Fitbit cloud storages.	Between January 2019 and December 2020, in Norway. Most of data collection was retrospective.
Data collection	Between April 2019 and February 2024, in San Diego, USA. Data collection was prospective.	Between April 2019 and September 2022, in the Moravian-Silesian and South Bohemian Regions of Czech Republic. Data collection was prospective.	Between September 2019 and October 2020, in Catalonia (Spain). Data collection was retrospective.	Between January 2019 and December 2020, in Norway. Most of data collection was retrospective.	Between January 2019 and December 2020, in Norway. Most of data collection was retrospective.
Ethical approval	Study approved by the Human Research Protections Program at UC San Diego (protocol #181862).	Study approved by the Ethics Committee of the University of Ostrava, Czech Republic.	Study approved by the Parc de Salut Mar Ethics Committee (GEIM-PS MAR, no. 2020/9307/I) and the Hospital Universitari Germans Trias i Pujol Ethics Committee (CEI no. PI20-182)	Study reviewed by The Regional Committees for Medical and Health Research Ethics North (reference 164780) and the Norwegian Centre for Research Data (reference 628485)	Study reviewed by The Regional Committees for Medical and Health Research Ethics North (reference 164780) and the Norwegian Centre for Research Data (reference 628485)
Date of the first COVID-19 lockdown	March 19th, 2020	March 16th, 2020	March 15th, 2020	March 12th, 2020	March 12th, 2020
Original sample size (included in this study)	638 (77)	1314 (65)	240 (48)	113 (36)	113 (36)
Activity monitor	Fitbit Charge 3	Fitbit Charge 3 and 4	Fitbit, Garmin, Withings, GoogleFit, Nokia, or Oura	Fitbit (35%) Garmin (65%)	Fitbit (35%) Garmin (65%)
Demographics of the participants included in the study^a	Biological sex: 45 (58%) Female Age: 24 (5)	Gender: 33 (51%) Female Age: 39 (9)	Gender: 24 (50%) Female Age: 53 (7)	Gender: 59 (56%) Female Age: 41 (11)	Gender: 59 (56%) Female Age: 41 (11)
Wear time information	Available	Available	Not available	Available only for Fitbit users	Available only for Fitbit users
Definition of valid days for step count data	At least 600 minutes of valid minute heart rate signal per day	At least 600 minutes of valid minute heart rate signal per day	At least one recorded step count. ^b	Fitbit users: At least 600 minutes of valid minute heart rate signal per day Garmin users: At least one recorded step count.	Fitbit users: At least 600 minutes of valid minute heart rate signal per day Garmin users: At least one recorded step count.

^aN (%); Mean (SD). For the open dataset, individual participant characteristics could not be linked to their respective physical activity data. Therefore, summary statistics are reported for the whole sample.

^bWe did not apply a higher step count threshold (e.g., 150 daily steps) because, in cases where both daily step counts and wear time data were available (e.g., Smart2.0 and 4HAIE studies), there were instances where wear time exceeded 10 hours, but daily steps ranged between 1–150. To account for such scenarios, we set the threshold to zero. This approach accounts for potential days spent in bed, such as during illness (e.g., COVID-19).

COVICAT (Catalunya, Spain). Catalunya entered its first COVID-19 state of emergency on March 14, 2020, enforcing strict home confinement until May 7, 2020. This was followed by a deconfinement period until June 21, 2020, during which various mobility restrictions, such as night curfews, city- or province-wide confinement, physical distancing, mask mandates, and leisure center closures, were imposed. A second state of emergency was declared on October 15, 2020.

Open dataset (Norway). The COVID-19 lockdown in Norway began on March 12, 2020, with the closure of gyms, schools, universities, kindergartens, and similar facilities. Individuals were encouraged to minimize group interactions, although outdoor exercise remained permitted. Social restrictions were progressively eased during April and May but were reintroduced incrementally starting in August 2020²³.

Participants

Smart 2.0. Participants ($N = 638$) were young adults (aged 18–35 years) recruited from three universities and five community colleges in San Diego, USA. The inclusion criteria included, (1) overweight or obese ($25 \leq \text{BMI} < 40 \text{ kg/m}^2$); (2) available for a 24-month intervention, (3) affiliated with one of the target universities as a student, staff, or alumni, (4) willing and able to use social media, a smartphone, text messaging, and Fitbit devices and app, and (5) willing and able to engage in moderate to vigorous physical activity. The exclusion criteria included, (1) comorbidities of obesity or conditions that prohibit compliance with the study protocol, (2) a recent cardiovascular event, (3) currently being treated for malignancy and/or an eating disorder, (4) planning to have weight loss surgery or enroll in a weight loss program, (5) loss of more than 15 pounds within the past 3 months, and (6) pregnancy or planning pregnancy within 24 months.

4HAIE. A total of $N = 1314$ participants aged 18–65 years were recruited for this study. Of these, $n = 747$ (57%) were physically active runners and $n = 567$ inactive controls (established at baseline by self-report). The participants were split between two regions of the Czech Republic. A highly polluted industrial region (Moravian-Silesian Region): $n = 750$ ($n = 436$ runners, $n = 314$ inactives) and a low-pollution level region (South Bohemian Region): $n = 564$ ($n = 311$ runners, $n = 253$ inactives). Inclusion criteria included (i) age between 18 and 65, (ii) permanent residency in either Moravia Silesia or Southern Bohemia regions for the past 5 years and no plan to move within a year, (iii) owning a smartphone and having internet access, (iv) being a non-smoker. The runners were defined as (i) regularly running for more than 6 weeks with at least 10 kilometers per week, (ii) fulfilling the weekly WHO PA recommendations²⁹, (iii) planning to continue running for the next 12 months. The inactive group included individuals who were capable of normal PA including running (i.e., no physician-diagnosed restrictions), but not meeting the recommended levels of PA.

COVICAT. The COVICAT study ($N = 10,087$) is a population-based investigation aimed at assessing the health impact of the COVID-19 pandemic on adults in Catalonia, Spain. It was built upon five pre-existing cohort studies established before the outbreak and inclusion criteria differed between each of the included cohorts. A description of each cohort including their inclusion criteria can be found here³⁰. All COVICAT participants were asked to fill out questionnaires and provided blood samples between May and December 2020 as part of the first assessment. Of the 10,087 individuals participating in the first wave of COVICAT, 2801 gave consent to be recontacted for the donation of physical activity data collected through their privately owned physical activity tracker. Of those, 240 participants eventually shared their physical activity data via Thryve.

Open dataset. Participants who owned an activity tracker (or smart watch) from Fitbit or Garmin, and who agreed to share physical activity-

related data collected from these devices between January 2019 and December 2020, were eligible for inclusion. Data were collected from 113 participants who shared their physical activity data using privately owned activity trackers (or smart watches) from Garmin and Fitbit.

Measures

In both SMART 2.0 and 4HAIE, daily steps were collected with a Fitbit Charge 3 or 4 (Fitbit Inc., San Francisco, CA, USA). Participants were asked to wear the wrist-worn device continuously during the trial according to the manufacturer's recommendations (i.e., to wear the device on one's non-dominant arm and up to three finger widths above the wrist bone). In the other two samples, data were retrospectively collected using crowdsourcing, i.e., participants consented to share their data collected via wearable devices with the research teams. Specifically, data collection extraction was supported by Thryve in ref. 24 and mSpider tool in ref. 23. In ref. 23, 35% of the participants wore a Fitbit device while the remaining 65% were Garmin users. Of the 240 COVICAT participants who provided step count data measured through physical activity trackers²⁴, 70 (29%) used a Fitbit, 134 (56%) owned a Garmin watch, 35 (15%) provided step count via Google Fit, and 5 (2%) through a Nokia device.

Rate of change in the physical activity stressor-response process.

Following a recent tutorial on physical activity resilience⁷, the rate of change in step count after the stressor was operationalized and implemented as follows. First, we defined the normal range in daily step count as the median number of daily steps during the five and a half months preceding the lockdown. Second, to model the physical activity trajectory following the stressor, we applied concurrent growth models to the step count time series following the beginning of the lockdown. Specifically, we fit the same versions of Generalized Additive Models (GAMs) through the R package *mgcv*³¹ and selected the best-fitting growth model based on the Root Mean Squared Error (RMSE), which tells how close the model's fitted values were to the observed data points. GAMs are particularly suited for defining the growth model, as they allow for flexible modeling of various non-linear patterns characterizing the time series' trajectory. In this analysis, the model's flexibility was specified by setting the number of basis functions (the k parameter in the *gam* function) to 20. More practically, this setting allowed each time series to be modeled with up to 20 degrees of freedom, providing sufficient flexibility to capture fluctuations on the order of roughly every 10 days across the observation period (from lockdown enforcement to the end of September). We initially tested $k = 10$ (approximately one fluctuation every 20 days). This yielded results highly comparable to those reported above in the Results section, with a very similar cross-correlation value between local DC and rate of change (-0.23 vs. -0.22), but with the peak correlation shifted slightly (lag -15 vs. lag -11). We ultimately selected $k = 20$, as this specification provided greater sensitivity to shorter-term fluctuations in step count trajectories and resulted in a more proximal alignment between local DC and subsequent changes in step count.

We extended this recently developed method⁷ by quantifying the rate of change in the stressor-response trajectory using the first-order derivative ($\frac{dy}{dt}$) of the growth function. The first-order derivative provides insight into the instantaneous rate of change of the step count trajectory, defining the resilience process. Positive values ($\frac{dy}{dt} > 0$) indicate that the function is increasing at that point, suggesting improvement or recovery. Negative values ($\frac{dy}{dt} < 0$) indicate that the function is decreasing, suggesting a decline or setback. A derivative of zero ($\frac{dy}{dt} = 0$) indicates a stationary point where there is no instantaneous change, possibly signifying a moment of equilibrium. The analysis of these derivatives allows a precise understanding of the direction and momentary pace of the stressor-response process over time. The derivatives were extracted with the *gratia* R package³².

Local Dynamic Complexity. Dynamic complexity (DC) was assessed by analyzing fluctuations in the step time series. The DC score is calculated

as the product of two measures: a fluctuation measure, F , and a distribution parameter, D ^{18,33}. The fluctuation measure is sensitive to both the amplitude and frequency of changes within the time series, peaking when data dynamics vary greatly from one observation to the next, with large and regular fluctuations (such as between the minimum and maximum values of the outcome). The distribution parameter quantifies the deviation of data points from the range of possible values in the time series, increasing with the occurrence of irregular changes from one observation to the next¹⁸. We refer to ‘local’ DC because the dynamic complexity scores were computed locally using a 14-day moving window throughout the time series. This means that the DC value on day n was derived from the period spanning from day $n-14$ to day n . The 14-day window was chosen iteratively: although we initially tested a 7-day window as done in¹³, cross-correlation (see next section on Data Analysis) with the rate of change was stronger using 14 days, yielding greater predictive potential.

Pre-lockdown step count (baseline). Baseline step count was defined as the median number of daily steps recorded between October 1, 2019, and the date of the country-specific lockdown enforcement. Details on the validity criteria for the baseline period (i.e., minimum of three valid weeks) are provided in the Data Analysis section.

Data analysis

Starting from the initial sample as described in the ‘Participants’ section, we applied the following pre-processing steps:

1. **Identification of valid daily step count data.** When wear time information was available (Smart 2.0 study, 4HAIE study, and Fitbit users from the open dataset), valid days were defined as days with at least 10 hours of wear time³⁴ and at least one step count registered. When wear time information was unavailable (COVICAT study and Garmin users from the open dataset), valid days were defined as days with at least one step count registered. In the latter case, we decided not to apply a higher step count threshold (e.g., 150 daily steps as done in²³) because, in cases where both daily step counts and wear time data were available (e.g., Smart 2.0 and 4HAIE studies), we observed instances where wear time exceeded 10 hours, yet daily steps fell within the range of 1–150. Non-valid days were recoded as missing values.
2. **Recoding upper bound outliers.** Daily step count observations above 40,000 have been recoded to 40,000 steps, as this approach prevents extreme values from influencing the results while preserving the integrity of the data by retaining realistic upper limits for step count.

Participants were drawn from the parent studies and included in the current analysis if they satisfied the following criteria:

1. **Valid baseline period:** Participants had at least three valid weeks of data during the pre-pandemic baseline period (October 1, 2019, to the date of the country-specific lockdown enforcement). A valid week was defined, based on prior research³⁵, as including at least one weekend day and two weekdays. The three-week requirement was chosen as a compromise to ensure a sufficiently stable baseline characterization while maximizing the number of participants included. This conservative threshold exceeds the minimum number of days typically recommended to obtain reliable step count estimates with accelerometry (i.e., 4–6 days)^{36,37}.
2. **Valid home confinement period:** Participants had valid data for all weeks within the four-week period following the country-specific lockdown enforcement, during stricter home confinement. For this period, a valid week was defined as including at least 3 valid days. Weekdays and weekend days were not differentiated, as the home confinement removed typical distinctions between them; therefore, all days within a week were treated equally.
3. **Valid recovery period:** Participants had at least 70% of valid weeks during the recovery period (four weeks after the country-specific lockdown enforcement through September 30, 2020). A valid week was

defined as including at least one weekend day and two weekdays, consistent with prior research³⁵.

4. **Missing data threshold:** To ensure sufficient data for reliable imputation and modeling, only participants with less than 20% missing values in their post-lockdown time series (from the date of the country-specific lockdown enforcement to September 30, 2020) were included.
5. **Step count perturbation:** After modeling growth functions (as described in the measures section), participants were included only if they experienced a perturbation of at least a 10% decrease in step counts during the 30 days following the lockdown relative to their median pre-lockdown step count level. This additional inclusion criterion is based on the study’s research question, which necessitates a perturbation in the system following the occurrence of the stressor.

Descriptive statistics of missing values, visualizations of their distribution, and trends in step counts during the post-stressor phase for the final analytical sample are presented in the supplemental materials (Supplementary Table 7, Supplementary Figs. 2–5).

The rate of change in the physical activity stressor-response process (dependent variable) was calculated as described in the measures section. Missing value imputation was not applied directly prior to modeling, as the growth model’s predicted values inherently provided estimates for originally missing observations. The threshold for defining a return to pre-stressor levels was operationalized as the median pre-lockdown step count, multiplied by a factor of 1.10 to account for the typically higher physical activity volumes in spring and summer compared to fall and winter^{38,39}.

Local dynamic complexity (independent variable) was calculated with an algorithm specifically designed to detect critical fluctuations in time series, implemented through the R package *casnet*⁴⁰. The function used to calculate the local DC metric does not permit missing values in the input time series. This is an important constraint, as missing values could result in an invalid metric, particularly for the distribution parameter, due to unequally spaced observations. To address this, missing daily step values were imputed using the Kalman Filter method, which is recommended for univariate time series imputation⁴¹. After data exploration and visual inspection of the cross-correlation plot for local DC calculated using 7-day and 14-day moving windows, we selected the 14-day window as it provided a more interpretable solution.

Before testing our main hypotheses, we examined the cross-correlation between local DC and the rate of change in the step count trajectory following the stressor. Cross-correlation is a statistical measure that determines the relationship between two time series, accounting for potential time lags. It is useful for identifying time delays and understanding how one series influences another. In our case, it aimed to identify the time lag at which local DC had its largest effect on the rate of change.

After the preliminary examination of the cross-correlations, we tested our hypotheses using a multilevel approach, implemented through the R package *lme4*⁴². The model included fixed effects for time (i.e., the progressive number of days since the start of the lockdown), local DC, and the quadratic term of local DC. Individual differences in the rate of change in the step count trajectory were accounted for by including a random intercept for the participant identifier. Adopting a ‘maximal approach’⁴³, random slopes were also estimated for time, local DC, and its quadratic term. The selection of these predictors, as both fixed and random effects, was pre-specified before running the analyses and guided by prior work¹³. Although the quadratic term for local DC was not included in that earlier study, we added it here because the findings suggested that the association between local DC and sudden changes in physical activity may follow a non-linear pattern: specifically, local DC was positively related to sudden losses but showed a weaker negative, non-significant associations with sudden gains. To assess this extension, we compared our final model with an otherwise identical model that excluded the quadratic term (both as a fixed and random effect). Results based on the Bayesian Information Criterion (BIC) indicated a better fit for the model, including the quadratic term (BIC decreased from 467,878 to 466,562). We focused on BIC as our performance metric because

it favors more parsimonious models compared to other criteria, such as the Akaike Information Criterion (AIC)⁴⁴. The linear and quadratic terms for local DC were taken from the point in time associated with the largest cross-correlation value. Finally, the pre-lockdown median number of steps (i.e., baseline steps) was included as a between-person variable to account for individual differences in physical activity level, under the assumption that more active participants would recover faster. The linear and quadratic terms for local DC were standardized, while time was scaled but not centered.

The multilevel equation, in the *lme4* syntax, is specified as follows:

$$\text{Rate of change} \sim \text{PreLockdown steps} + \text{Time} + \text{LDC} + (\text{LDC})^2 + (\text{Time} + \text{LDC} + (\text{LDC})^2 | \text{Participant ID})$$

For the sensitivity analysis, we ran four additional models—one for each data source (Smart 2.0, 4HAIE, COVICAT, and the Open dataset)—to test whether the effect of local DC on the rate of change was moderated by the study data source (operationalized as a dummy variable). We also conducted a sensitivity analysis that included only complete and partially complete cases. In addition, we examined whether device brand moderated the effect of local DC on the rate of change. Finally, we tested a model including age as a predictor. For this latter analysis, participants from the Open dataset were excluded, as person-specific age information was not available. The results of these sensitivity analyses are presented in the Supplementary Tables 2–6.

Data availability

The data used in the analysis are available at <https://osf.io/gsmhk/>.

Code availability

The R scripts used for the analysis are available at <https://osf.io/gsmhk/>.

Received: 11 December 2024; Accepted: 8 September 2025;

Published online: 09 October 2025

References

- Warburton, D. E. R. & Bredin, S. S. D. Health benefits of physical activity: a systematic review of current systematic reviews. *Curr. Opin. Cardiol.* **32**, 541–556 (2017).
- Scheffer, M. et al. Quantifying resilience of humans and other animals. *Proc. Natl Acad. Sci. USA* **115**, 11883–11890 (2018).
- Hill, Y., Kiefer, A. W., Oudejans, R. R. D., Baetzner, A. S. & Den Hartigh, R. J. R. Adaptation to stressors: Hormesis as a framework for human performance. *N. Ideas Psychol.* **73**, 101073 (2024).
- Spence, J. C. & Lee, R. E. Toward a comprehensive model of physical activity. *Psychol. Sport Exerc.* **4**, 7–24 (2003).
- Den Hartigh, R. J. R. & Hill, Y. Conceptualizing and measuring psychological resilience: What can we learn from physics?. *N. Ideas Psychol.* **66**, 100934 (2022).
- Carver, C. S. Resilience and thriving: issues, models, and linkages. *J. Soc. Issues* **54**, 245–266 (1998).
- Baretta, D. et al. Resilience characterized and quantified from physical activity data: a tutorial in R. *Psychol. Sport Exerc.* **65**, 102361 (2023).
- Nahum-Shani, I. et al. Just-in-Time Adaptive Interventions (JITIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann. Behav. Med.* **52**, 446–462 (2018).
- Heino, M. T. J., Proverbio, D., Marchand, G., Resnicow, K. & Hankonen, N. Attractor landscapes: a unifying conceptual model for understanding behaviour change across scales of observation. *Health Psychol. Rev.* 1–18 <https://doi.org/10.1080/17437199.2022.2146598> (2022).
- Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E. H. & Rietkerk, M. Early warning signals also precede non-catastrophic transitions. *Oikos* **122**, 641–648 (2013).
- Scheffer, M. et al. Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
- Helmich, M. A. et al. Early warning signals and critical transitions in psychopathology: challenges and recommendations. *Curr. Opin. Psychol.* **41**, 51–58 (2021).
- Chevance, G. et al. Characterizing and predicting person-specific, day-to-day, fluctuations in walking behavior. *PLoS ONE* **16**, e0251659 (2021).
- Baretta, D., Mazéas, A., Chalabaev, A., Inauen, J. & Chevance, G. Early warning signals of sudden changes in walking behavior. https://doi.org/10.31219/osf.io/wpge7_v1 (2025).
- Wu, Z. et al. The COVID-19 pandemic and daily steps in the general population: meta-analysis of observational studies. *JMIR Public Health Surveill.* **9**, e40650 (2023).
- Paluch, A. E. et al. Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. *Lancet Public Health* **7**, e219–e228 (2022).
- Lee, I.-M., Keadle, S. K. & Matthews, C. E. Fitness trackers to guide advice on activity prescription. *JAMA* **330**, 1733–1734 (2023).
- Schiepek, G. & Strunk, G. The identification of critical fluctuations and phase transitions in short term and coarse-grained time series—a method for the real-time monitoring of human change processes. *Biol. Cyber.* **102**, 197–207 (2010).
- Evenson, K. R. & Spade, C. L. Review of Validity and Reliability of Garmin Activity Trackers. *J. Meas. Phys. Behav.* **3**, 170–185 (2020).
- Feehan, L. M. et al. Accuracy of fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth* **6**, e10527 (2018).
- Tedesco, S. et al. Validity evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in free-living environments in an older adult cohort. *JMIR mHealth uHealth* **7**, e13084 (2019).
- Elavsky, S. et al. Physical activity in an air-polluted environment: behavioral, psychological and neuroimaging protocol for a prospective cohort study (Healthy Aging in Industrial Environment study—Program 4). *BMC Public Health* **21**, 126 (2021).
- Henriksen, A., Johannessen, E., Hartvigsen, G., Grimsgaard, S. & Hopstock, L. A. Dataset of consumer-based activity trackers as a tool for physical activity monitoring in epidemiological studies during the COVID-19 Pandemic. *Data Brief.* **41**, 108003 (2022).
- kogevinas, M. et al. Ambient air pollution in relation to SARS-CoV-2 infection, antibody response, and COVID-19 disease: a cohort study in Catalonia, Spain (COVICAT Study). *Environ. Health Perspect.* **129**, 117003 (2021).
- Jin, D. et al. Self-tracking behaviour in physical activity: a systematic review of drivers and outcomes of fitness tracking. *Behav. Inf. Technol.* **41**, 242–261 (2022).
- Mansour-Assi, S. J. et al. Social Mobile Approaches to Reducing Weight (SMART) 2.0: protocol of a randomized controlled trial among young adults in university settings. *Trials* **23**, 7 (2022).
- Costello, V. L. et al. Impact of the COVID-19 pandemic on objectively measured physical activity and sedentary behavior among overweight young adults: yearlong longitudinal analysis. *JMIR Public Health Surveill.* **7**, e28317 (2021).
- Vládní usnesení související s bojem proti epidemii—rok 2020. <https://vlada.gov.cz/epidemie-koronaviru/dulezite-informace/vladni-usneseni-souvisejici-s-bojem-proti-epidemii-koronaviru-rok-2020-186999/>.
- Bull, F. C. et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* **54**, 1451–1462 (2020).
- Goldberg, X. et al. Mental health and COVID-19 in a general population cohort in Spain (COVICAT study). *Soc. Psychiatry Psychiatr. Epidemiol.* **57**, 2457–2468 (2022).
- Wood, S. N. *Generalized Additive Models: An Introduction with R*. (Chapman and Hall/CRC, 2017). <https://doi.org/10.1201/9781315370279>

32. Simpson, G. L. gratia: An R package for exploring generalized additive models. *J. Open Source Softw* **9**, 6962, <https://doi.org/10.21105/joss.06962> (2024).
 33. Olthof, M. et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clin. Psychol. Sci.* **8**, 25–35 (2020).
 34. Matthews, C. E., Hagströmer, M., Pober, D. M. & Bowles, H. R. Best practices for using physical activity monitors in population-based research. *Med Sci. Sports Exerc* **44**, S68–S76 (2012).
 35. Tudor-Locke, C. et al. How many days of pedometer monitoring predict weekly physical activity in adults?. *Preventive Med.* **40**, 293–298 (2005).
 36. Hart, T. L., Swartz, A. M., Cashin, S. E. & Strath, S. J. How many days of monitoring predict physical activity and sedentary behaviour in older adults?. *Int J. Behav. Nutr. Phys. Act.* **8**, 62 (2011).
 37. Kang, M. et al. How many days are enough? a study of 365 days of pedometer monitoring. *Res. Q. Exerc. Sport* **80**, 445–453 (2009).
 38. Turrisi, T. B. et al. Seasons, weather, and device-measured movement behaviors: a scoping review from 2006 to 2020. *Int. J. Behav. Nutr. Phys. Act.* **18**, 24 (2021).
 39. Clemes, S. A., Hamilton, S. L. & Griffiths, P. L. Summer to winter variability in the step counts of normal weight and overweight adults living in the UK. *J. Phys. Act. Health* **8**, 36–44 (2011).
 40. Hasselman, F., Olthof, M. & Cui, J. casnet: A toolbox for studying Complex Adaptive Systems and NETworks (0.2.2), <https://github.com/FredHasselman/casnet> (2022).
 41. Gómez, Maravall, V. & Estimation, A. Prediction, and Interpolation for Nonstationary Series with the Kalman Filter. *J. Am. Stat. Assoc.* **89**, 611–624 (1994).
 42. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
 43. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
 44. Chakrabarti, A. & Ghosh, J. K. in *Philosophy of Statistics* (eds. Bandyopadhyay, P. S. & Forster, M. R.) 7 583–605 (North-Holland, Amsterdam, 2011).
 45. Obón-Santacana, M. et al. GCAT|Genomes for life: a prospective cohort study of the genomes of Catalonia. *BMJ Open* **8**, e018324 (2018).
- Rapid Response activity 20873A and SR20-01024 La Caixa Foundation. The Barcelona Institute for Global Health (ISGlobal) receives support from the grant CEX2023-0001290-S funded by MCIN/AEI/10.13039/501100011033, and support from the Generalitat de Catalunya through the CERCA Program. GC has also been awarded with the grant RYC2021-033537-I, supported by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR"; and from the Ministry of Research and Universities of the Government of Catalonia (2021 SGR 01563). The author SK and this project were supported by the European Commission through a MSCA-IF (10.3030/840513). The author JB received funding from the Juan de la Cierva Formación (FJC2021-046458-I), financed by MICIU/AEI/10.13039/501100011033 and by the European Union Next-GenerationEU/PRTR.

Author contributions

D.B. and G.C. conceptualized the study and defined the methodology. A.L.A., M.O. and R.d.H. provided theoretical guidance. J.G., S.K. and S.E. acquired funding for the projects that provided the data for this study. S.E., L.K., S.K., J.G.A. and J.G. administered the projects that provided the data for this study. D.B. conducted the statistical analysis and developed the R code. J.B. provided methodological feedback. D.B. and G.C. wrote the original draft of the manuscript. All authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-01998-0>.

Correspondence and requests for materials should be addressed to Dario Baretta.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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 Author(s) 2025

Acknowledgements

The manuscript has been prepared with the support from the project "Research of Excellence on Digital Technologies and Wellbeing CZ.02.01.01/00/22_008/0004583" which is co-financed by the European Union. The manuscript has also been prepared with the support of the European Union under the LERCO project (CZ.10.03.01/00/22_003/0000003) via the Operational Programme Just Transition. We would like to thank the project "Healthy Aging in Industrial Environment HAIE CZ.02.1.01/0.0/0.0/16_019/0000798" which was co-financed by the European Union for the use of the data. The SMART2.0 study was supported by a grant from the National Institute of Health's National Heart Lung and Blood Institute: R01HL136769. The COVICAT data used in this study were generated by the GCAT-Genomes for Life Cohort study of the Genomes of Catalonia, Fundació IGTP and COVICAT Study Group (Cohort Covid de Catalunya). COVICAT is supported by ISGlobal and IGTP and funded by EIT COVID-19