

A physiological heart-rate model for energy expenditure across work, exercise, and recovery: development and in-sample evaluation against indirect calorimetry

Agnar Steinarsson

University of Iceland

Jón Hlöðver Friðriksson

Driftline

Gréta Jakobsdóttir

University of Iceland

Guillaume Bonnier

Swiss Center for Electronics and Microtechnology (Switzerland)

Christopher Moufawad El Achkar

Swiss Center for Electronics and Microtechnology (Switzerland)

Erlingur Jóhannsson

erljo@hi.is

University of Iceland

Alia Lemkaddem

Swiss Center for Electronics and Microtechnology (Switzerland)

Research Article

Keywords: energy expenditure, heart rate, indirect calorimetry, wearable sensors, model development

Posted Date: October 26th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11319899/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: Competing interest reported. AS, the PhD candidate supervised by EJ and GJ, is the founder of Driftline, a company developing heart-rate-based fitness and metabolic tracking technologies. AS is the inventor of the physiological heart-rate model described in this study. A PCT patent application covering aspects of this method has been filed. JHF was employed by Driftline at the time of the study. The remaining authors declare no commercial or financial relationships that could be construed as a potential conflict of interest. The study was conducted under academic supervision (EJ and GJ) and approved by the relevant institutional bioethics committee.

Abstract

Purpose To develop and evaluate a physiological heart-rate model (PHM) for energy expenditure (EE) across heterogeneous activity states, and to examine total-session agreement, state-dependent estimation error, and estimation of sedentary and recovery EE without age-, sex-, or height-based population equations. We developed and evaluated PHM using heart rate as the sole continuous physiological signal.

Methods Twenty healthy adults (11 men, 9 women; age 39.5 ± 7.3 years) completed office work, a non-rhythmic exercise circuit, graded treadmill exercise to exhaustion, and seated recovery while EE was measured by indirect calorimetry (COSMED K5). The PHM used heart rate, maximal heart rate (HR_{max}), speed at the second ventilatory threshold (vVT_2), an endurance parameter (E), body mass, and protocol-derived activity-state labels. Empirical coefficients were calibrated in this cohort; performance therefore represents in-sample evaluation.

Mean session EE was 465.9 versus 466.8 kcal by indirect calorimetry. Bland-Altman bias was -0.9 kcal (95% limits of agreement -33.6 to +31.7 kcal), and Lin's concordance correlation was 0.98. Energy-weighted mean absolute deviation (MAD) across conditions was 8.7%, lowest during running (3.5%) and largest during stair descent (35.3%) and weighted exercise (21.2%). Sedentary work showed +0.6% bias (individual SD 11.2%); recovery was overestimated by 15.7%.

The PHM showed promising in-sample performance across diverse activity states without age-, sex-, or height-based population equations, although individual sedentary error resembled that of population equations and individual inputs came from the evaluated session. External validation with coefficients fixed a priori and evaluation of automatic activity-state detection are required.

Introduction

Energy expenditure (EE) is fundamental to exercise science, weight management, and metabolic health, and wearable technology has made continuous estimation commonplace. However, accurate estimation remains difficult across the heterogeneous activities of daily life. Most consumer and research devices estimate EE primarily from accelerometry, alone or with physiological sensors. Because movement and metabolic demand are activity-dependent, static and resistance exercise, cycling, load carriage, and recovery can impose substantial metabolic cost with little acceleration (Chen and Bassett 2005). Accelerometer equations and activity-classification pipelines therefore perform inconsistently across everyday activities (Lyden et al. 2011; Shcherbina et al. 2017).

Heart rate offers a complementary signal because it reflects the cardiovascular response to metabolic demand and is continuously available from wearables (Achten and Jeukendrup 2003). Heart-rate methods can estimate oxygen uptake and EE reproducibly under standardized conditions (Wicks et al. 2011; Schantz et al. 2019), but flex-heart-rate approaches require individual laboratory calibration, while population-based models show larger individual error, particularly at low intensity (Spurr et al. 1988; Rennie et al. 2001; Keytel et al. 2005). Recovery and non-rhythmic or resistance effort are especially problematic because heart rate may remain elevated or rise out of proportion to whole-body oxygen uptake (Achten and Jeukendrup 2003; Børsheim and Bahr 2003).

We developed a physiological heart-rate model (PHM) that individualizes the heart-rate-metabolic-demand relationship using maximal heart rate (HR_{max}), speed at the second ventilatory threshold (vVT_2), an endurance parameter (E), body mass, and quantities from a heart-rate-speed framework (Steinarsson et al., unpublished manuscript). Heart rate is the sole continuous physiological input; activity-state labels select predefined parameters for work/other activity, walking, exercise, and recovery. Such states could potentially be identified automatically from accelerometry, although this was not evaluated here. The model uses no age-, sex-, or height-based population equation for EE or resting metabolism.

The primary aim of this study was to develop the PHM and characterize its apparent in-sample agreement with indirect calorimetry across a heterogeneous protocol spanning sedentary work, non-rhythmic and resistance exercise, graded treadmill walking and running, and post-exercise recovery. Specifically, we addressed three research questions: (1) how closely does PHM reproduce total session energy expenditure; (2) how does estimation accuracy vary across protocol phases and individual activity states, particularly locomotion, resistance/eccentric exercise, and recovery; and (3) can PHM estimate light sedentary and recovery energy expenditure without relying on age-, sex-, or height-based population equations?

Materials and methods

Study design and participants

The study was part of the METFIT PhD study at the University of Iceland and an integral part of the FITSILVER project (Eurostars-3 Project E! 2671), a collaboration between Driftline (Iceland) and CSEM (Switzerland) from 2023 to 2025. It was performed at the CSEM Sports Science Laboratory in Neuchâtel, Switzerland, approved by the Swiss Ethics CER-VD (BASEC ID 2024 – 01496), and conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before participating.

Twenty-eight healthy adults were enrolled in the study; eight were excluded before model evaluation because of insufficient data quality (heart-rate signal dropout or incomplete gas-exchange recording in one or more phases), leaving 20 for analysis (11 men, 9 women; age 39.5 ± 7.3 years; body-mass index 22.8 ± 2.0 kg·m⁻²). Participants were recreationally active, free of cardiovascular, metabolic, and musculoskeletal disease, and were not taking chronotropic medication. Each participant completed a single recording session comprising three consecutive phases performed on a common, continuously recorded timeline (Fig. 1). Phase A (Work) consisted of 50–60 min of seated and standing office work activity. Phase B (Exercise) was a circuit of non-rhythmic activities: stationary cycling (3 min), bicep curls (30 s), lunges (60 s), push-ups (30 s), squats (30 s), planks (60 s), stair ascent (30 s), and stair descent (30 s), with 60 s rest intervals between activities, with a total time of 15 min. External loads and intensities were not standardized; participants performed the tasks at self-selected moderate effort.

Phase C (CPET + recovery) began with a standardized preparation period of 3 min lying, 3 min sitting, and 3 min standing, followed by an incremental treadmill protocol at 0% gradient that started at 1.8 km·h⁻¹ and increased by 1.8 km·h⁻¹ every 3 min. Participants walked through the lower stages and transitioned to running at a self-selected speed, continuing to volitional exhaustion, after which they completed a 20-min seated recovery. The onset and offset of every activity mode were annotated on the shared timeline and used both to segment subsequent analyses and to supply the broad activity-state labels required by the PHM.

Reference energy expenditure and heart rate

Energy expenditure (EE) was measured breath-by-breath by indirect calorimetry (COSMED K5; COSMED, Rome, Italy) (Guidetti et al. 2018; Perez-Suarez et al. 2018) throughout the protocol. The analyzer was calibrated before each session against reference and ambient gases and with a 3-L syringe for the flowmeter. Reference EE (EE_{ref}) was computed from oxygen uptake and carbon-dioxide production using the Weir equation (Weir 1949) and reduced to a common 1-s resolution. EE_{ref} served as the criterion measure throughout. Heart rate was recorded continuously at 1-s resolution using a Polar H10 chest-strap sensor (Polar Electro Oy, Kempele, Finland). Heart rate and reference EE traces were time-aligned before analysis.

Determination of model inputs

Maximal heart rate ($HR_{\max}$) and the running speed at the second ventilatory threshold (vVT_2) were obtained from the graded treadmill test. $HR_{\max}$ was set 0–4 bpm above the highest heart rate recorded during the CPET, because peak test heart rate typically falls slightly short of true maximum (by 3.8 bpm on average in the companion study; Steinarsson et al., unpublished manuscript); the value within this range was chosen during model fitting. $VO_{2\max}$ was taken as the highest 15-s averaging interval, and VT_2 was identified as the respiratory compensation point using the ventilatory-equivalents method (Binder et al. 2008; Keir et al. 2024). The endurance parameter E was obtained from each participant's heart-rate response during recovery from the CPET using the heart-rate–speed framework described in Online Resource 1. All three individual inputs therefore came from the session used for evaluation (see Limitations).

Within this framework, E and $HR_{\max}$ define the heart-rate threshold interval P_i . P_i is numerically equivalent to the model-predicted resting heart rate but was not obtained from an independent resting-HR measurement in the present study. When E is unavailable, P_i may alternatively be supplied from a measured or independently estimated resting heart rate. Likewise, E may be estimated independently from the ratio between second threshold and maximum running speed using the E -ratio relationship described in the companion study (Steinarsson et al., unpublished manuscript). These represent alternative parameterization routes rather than additional simultaneous inputs.

Model development and calibration

Empirical coefficients identified as [emp] in Online Resource 1, together with the P_c multiplier and the $HR_{\max}$ offset, were fine-tuned using the present dataset, with no held-out or resampling-based validation. Protocol-derived activity-state labels selected predefined parameters for work/other activity, walking, exercise, and recovery. Thus, all reported performance is apparent in-sample performance; the final coefficients should be fixed a priori for external validation.

The physiological heart-rate model

The PHM estimates EE each second from heart rate using $HR_{\max}$, vVT_2 , E , body mass, and quantities from the heart-rate–speed framework (Online Resource 1). Heart rate is the sole continuous physiological signal; activity-state labels select parameters but contribute no quantitative movement data. In the companion study, E fitted from submaximal running and recovery heart rate was evaluated against a CPET-based reference and was closely related to the ratio of maximal to threshold running speed (Steinarsson et al., unpublished manuscript). Here, E was fitted from recovery after the CPET. Exercise-circuit EE was scaled by ϕ (a cohort-tuned constant, 2.1; not applied to treadmill walking or running), and piecewise corrections adjusted work and recovery multipliers for atypical heart-rate behavior when mean work heart rate exceeded $1.4 \cdot P_i$ or early-recovery heart rate exceeded late-recovery heart rate by more than 15%; these corrections were applied in 2 of 20 participants (Online Resource 1, S1.7).

Framework-derived quantities

From $HR_{\max}$ and E , the heart-rate–speed framework defines the heart-rate threshold interval P_i , from which the threshold heart rates are derived. The resting/response anchor was set empirically to $P_c = 1.16 \cdot P_i$. This differs from the framework anchor used in the companion study, $P_c = c_{\text{rest}} \cdot P_i$ with $c_{\text{rest}} = 2(4 + E)/(8 - E)$ (≈ 1.20 at the cohort mean E). The value 1.16 was chosen because it minimized session bias in this cohort; using the framework anchor, with E refitted, would lower Work-phase EE by approximately 3% and running EE by less than 1% (Online Resource 1, S1.2). In the model geometry, P_i corresponds numerically to predicted resting heart rate, although measured resting heart rate was not used to derive P_i in the present analysis. Together with vVT_2 , the framework also provides maximal aerobic speed (MAS) and maximal

aerobic power (MAP). Body mass enters the calculations through MAP. Cardiovascular drift is modelled separately as a time-varying component of the heart-rate response (Online Resource 1, S1.4).

Metabolic calibration

The activity factor AF ($\text{beats} \cdot \text{kcal}^{-1}$), representing the cardiac cost per unit of energy at high exercise intensity (Swain & Leutholtz 1997), was derived from heart-rate reserve and aerobic power:

$$AF = 1.7 + 16.1 \cdot (HR_{max} - P_i) / MAP$$

Basal metabolic rate (BMR, $\text{kcal} \cdot \text{min}^{-1}$) was obtained from the activity factor rather than from a generic age-, sex-, or height-based anthropometric equation:

$$BMR = (2554.6 - 92.9 \cdot AF) / 1440$$

The low-intensity basal factor BF ($\text{beats} \cdot \text{kcal}^{-1}$) was defined as $BF = P_c / BMR$. Calibration of these empirical relationships is detailed in Online Resource 1.

Energy-expenditure estimation

For each second t , EE was estimated as a resting baseline plus a heart-rate-driven term:

$$EE(t) = c_{rest} \cdot BMR + [HR(t) - P_c] / \alpha(t)$$

where BMR is the basal metabolic rate ($\text{kcal} \cdot \text{min}^{-1}$), c_{rest} an E -dependent resting multiplier, and P_c the resting heart-rate anchor. The instantaneous activity factor $\alpha(t)$ rises from its low-intensity value BF to its high-intensity value AF through a logistic transition in the metabolic heart rate $HR_m(t)$, with an added drift term above the VT_2 heart rate (P_2):

$$\alpha(t) = AF + (BF - AF) / (1 + \exp\{k \cdot [HR_m(t) - HR_c(t)]\}) - \delta \cdot \max(0, HR_m(t) - P_2)$$

Here k sets transition steepness and δ the supra-threshold drift. The reference heart rate $HR_c(t) = m \cdot P_c$ is state-specific, with predefined multipliers for work, exercise, recovery, and walking. $HR_m(t)$ is measured heart rate bounded below by P_c and, during recovery, relaxed toward P_c at rate r . Exercise-phase EE was scaled by φ , and piecewise corrections adjusted work and recovery multipliers for atypical heart-rate behavior; these corrections were applied in two of 20 participants (Online Resource 1).

EE estimates were summed within each activity mode to yield mode-level energy (kcal); modes were then aggregated into phases (A, B, C) and into the whole session. Estimated and reference EE were binned on the same annotated protocol timeline so that both series were compared over identical time windows for every mode, phase, and the session total.

Outcome measures

Outcome measures

Performance was quantified at three levels. For each activity mode and phase, signed relative deviation and mean absolute deviation (MAD) were calculated from segment energy as $(EE_{est} - EE_{ref}) / EE_{ref}$. For each phase and the whole session, absolute cumulative deviation quantified total-energy error, while energy-weighted MAD summarized per-mode error without cross-mode cancellation. Whole-session agreement was also evaluated from participant session

totals. Second-by-second percentage error was not used because breath-by-breath variability in the criterion measure inflates it independently of model performance.

Statistical analysis

Data are presented as mean $\pm$ SD unless otherwise stated. Mode- and phase-level deviations were summarized as above; whole-session totals were assessed with Bland-Altman bias and 95% limits of agreement, root-mean-square error, Pearson correlation, Lin's concordance correlation coefficient, and ordinary-least-squares regression (Bland and Altman 1986; Lin 1989). Sex differences (Online Resource 1, Table S4) were explored with independent-samples *t* tests or Mann-Whitney U tests when Shapiro-Wilk indicated non-normality; these small-group comparisons were descriptive and uncorrected for multiplicity. Significance was set at $\alpha = 0.05$. Analyses used jamovi v2.6.44 (The jamovi project 2025). Because coefficients were calibrated in the same cohort, all metrics describe apparent in-sample performance.

Generative AI statement

Claude (Anthropic) and ChatGPT (OpenAI) assisted with manuscript drafting, organization, and language revision. Claude also assisted with figure preparation using author-supplied data. All outputs were critically reviewed and verified by the authors, who approved the final manuscript and take full responsibility for its content.

Results

Descriptive statistics

The participant characteristics and main CPET results are summarized in Table 1.

Table 1 Descriptive statistics and main results from indirect calorimetry			
Characteristic	Men (<i>n</i> = 11)	Women (<i>n</i> = 9)	All (<i>n</i> = 20)
Age (years)	41.7 $\pm$ 6.1	36.8 $\pm$ 8.0	39.5 $\pm$ 7.3
Height (cm)	178 $\pm$ 4	165 $\pm$ 5	172 $\pm$ 8
Body mass (kg)	75.3 $\pm$ 6.0	58.2 $\pm$ 3.5	67.6 $\pm$ 10.0
BMI (kg·m ⁻²)	23.8 $\pm$ 2.0	21.5 $\pm$ 0.8	22.8 $\pm$ 2.0
VO _{2max} (mL·kg ⁻¹ ·min ⁻¹)	46.6 $\pm$ 5.8	43.5 $\pm$ 4.6	45.2 $\pm$ 5.4
HR _{max} (bpm)	188 $\pm$ 13	185 $\pm$ 12	187 $\pm$ 12
vVT ₂ (km·h ⁻¹)	11.8 $\pm$ 1.6	9.7 $\pm$ 1.2	10.9 $\pm$ 1.8

BMI, body mass index; VO_{2max}, maximal oxygen uptake; HR_{max}, maximal heart rate; vVT₂, running speed at the VT₂ threshold.

Whole-session agreement

To address the first research question, whole-session PHM estimates were compared with indirect calorimetry. Across the whole session, mean EE was 465.9 kcal for PHM and 466.8 kcal for indirect calorimetry (Fig. 2; Online Resource 1, Table S3). Bland-Altman bias was -0.9 kcal (-0.2% of mean reference EE), with 95% limits of agreement from -33.6 to $+31.7$ kcal (-7.2% to $+6.8\%$ of mean reference EE); RMSE was 16.3 kcal (3.5%). Session totals were strongly associated (Pearson $r = 0.98$; Lin's concordance = 0.98); OLS regression gave a slope of 0.92 and intercept of 37.7 kcal (Fig. 2B).

Mean signed participant-level session deviation was $+0.1 \pm 3.8\%$, and mean absolute cumulative deviation was $3.1 \pm 2.2\%$ (Online Resource 1, Tables S3 and S5).

Agreement by protocol phase

To address the second research question, estimation error was examined across protocol phases and individual activity states. Mean absolute cumulative deviation was 9.2% for Work, 6.8% for Exercise, and 2.7% for CPET/recovery, with signed deviations within $\pm 1\%$ (Table 2). Weighted MAD was 9.2%, 13.0%, and 7.7%, respectively. In Exercise, the lower cumulative deviation than weighted MAD (6.8% vs 13.0%) reflected cancellation of opposite-signed mode errors.

Table 2
Agreement by protocol phase and overall session ($n = 20$).

Level	Cumulative deviation, signed (%)	Cumulative deviation (%)	Weighted MAD (%)
Work	$+0.6 \pm 11.2$	9.2 ± 6.0	9.2 ± 6.0
Exercise	$+0.7 \pm 8.5$	6.8 ± 4.9	13.0 ± 2.8
CPET	-0.1 ± 3.2	2.7 ± 1.6	7.7 ± 2.5
Overall (session)	$+0.1 \pm 3.8$	3.1 ± 2.2	8.7 ± 1.9

Cumulative deviation = $(\Sigma\text{PHM} - \Sigma\text{ref})/\Sigma\text{ref}$ over the level (signed, and absolute). Weighted MAD = energy-weighted mean of per-mode |deviation| within the level. Cumulative |deviation| reflects total-energy error; weighted MAD reflects across-condition error and excludes cross-mode cancellation.

Agreement by activity mode

Mode-level analysis further addressed the second research question. In-sample performance varied across the fourteen activity modes (Table 3, Fig. 3). Running had the lowest error (bias -3.1% , MAD 3.5%) and walking remained relatively close (bias -5.9% , MAD 10.8%). Errors were larger in brief non-rhythmic and eccentric modes: stair descent showed $+35.3\%$ bias and weighted exercise $+18.1\%$ bias (MAD 21.2%), while push-ups (-9.2%) and squats (-4.3%) were underestimated. To address the third research question, light sedentary work and recovery were examined specifically. Sedentary work showed $+0.6\%$ bias and 9.2% MAD; the participant-level SD of 11.2% corresponds to approximate 95% limits of -21% to $+22\%$, and 10 of 20 participants were within $\pm 10\%$ sedentary work bias. Recovery EE was systematically overestimated by 15.7%.

The largest percentage errors occurred in low-energy modes (approximately 4–12 kcal), limiting their contribution to session totals. Apparent session performance was similar for men and women (Online Resource 1, Table S4).

Table 3
Agreement between PHM-estimated and reference energy expenditure by activity mode ($n = 20$).

Activity mode	Phase	Reference EE (kcal)	Bias (%)	MAD (%)
Work	Work	108.2 ± 15.3	+ 0.6 ± 11.2	9.2 ± 6.0
Bike	Exercise	18.3 ± 4.2	-5.3 ± 10.0	10.1 ± 4.8
Weights	Exercise	5.8 ± 1.6	+ 18.1 ± 15.9	21.2 ± 11.2
Lunges	Exercise	7.0 ± 1.2	-3.9 ± 13.2	11.5 ± 7.1
Pushups	Exercise	7.0 ± 1.5	-9.2 ± 10.9	12.0 ± 7.5
Squats	Exercise	6.9 ± 1.9	-4.3 ± 11.6	9.9 ± 7.1
Planks	Exercise	8.7 ± 1.8	+ 9.1 ± 11.4	11.2 ± 9.2
Stairs (up)	Exercise	5.3 ± 1.5	-2.6 ± 12.7	9.7 ± 8.4
Stairs (down)	Exercise	3.5 ± 0.9	+ 35.3 ± 14.2	35.3 ± 14.2
Prep	CPET	12.1 ± 8.3	-10.4 ± 10.6	12.0 ± 8.7
Rest	CPET	22.2 ± 6.5	+ 5.4 ± 17.6	13.6 ± 12.1
Walk	CPET	49.4 ± 10.9	-5.9 ± 11.8	10.8 ± 7.3
Run	CPET	162.1 ± 47.8	-3.1 ± 2.7	3.5 ± 2.1
Recovery	CPET	50.3 ± 15.1	+ 15.7 ± 9.3	15.7 ± 9.3

EE, energy expenditure; MAD, mean absolute deviation. Reference EE is the mean ± SD of energy expended per mode (COSMED K5 indirect calorimetry). Deviation = (PHM - reference)/reference, computed per participant from energy summed over each mode; bias is the signed mean and MAD the mean absolute value, each reported as mean ± SD across participants.

Whole-session time-series agreement

Representative whole-session traces are shown in Fig. 4A. PHM-estimated EE tracked the reference through work, exercise, graded locomotion, and recovery. During recovery, both estimated and reference EE declined despite persistently elevated heart rate, although group-level recovery energy was overestimated. Figure 4B illustrates the cohort-mean exercise circuit, with visible overestimation during resistance and stair descent. On average the model overestimated the resistance and eccentric modes (Weights, Planks, Stairs-down; shaded) and underestimated the concentric bodyweight modes, consistent with the mode-level pattern in Fig. 3.

Discussion

The present study addressed three questions about the apparent in-sample performance of PHM. First, PHM reproduced total session EE closely, although errors were larger at the individual activity-mode level. Second, accuracy varied systematically among activity states and modes, with the closest agreement during rhythmic locomotion and larger errors during recovery and some resistance or eccentric activities. Third, light sedentary EE was estimated with little mean bias without explicit age-, sex-, or height-based population equations, although individual error was similar to that of population equations and the relevant relationships were calibrated in this cohort. These findings support the physiological basis of the model while identifying activity context and external validation as the main priorities for further development.

Total-session accuracy

PHM reproduced total session EE closely in this development cohort, answering research question 1. Mean session bias was small, absolute cumulative deviation was 3.1%, and concordance with indirect calorimetry was high. The larger energy-weighted per-mode MAD of 8.7% shows, however, that accurate session totals do not imply equally accurate estimates for every activity. Total-session error is necessarily weighted by the activities performed. Positive and negative mode-specific errors may partly offset one another, so a near-zero total bias can conceal larger errors within individual modes. A day dominated by sedentary activity, walking, and running would give more weight to modes that performed well here, whereas prolonged resistance exercise, cycling, or repeated recovery periods would change the error profile.

This distinction between good average agreement and greater individual uncertainty is also seen in free-living EE studies. Brage et al. (2015), using heart rate and movement sensing against doubly labelled water over 14 days, reported total-EE biases below 5%, while precision still depended on the sensing method and individual calibration. Motion-sensor estimates can perform reasonably at the population level while retaining substantial individual error, with additional physiological information such as heart rate improving PAEE estimation (Sardinha and Júdice 2017). The present session-level results are therefore encouraging, but free-living validation is needed before the observed aggregate accuracy can be generalized to daily EE.

Accuracy across activity states and modes

Accuracy differed among activity states and modes, answering research question 2. Running showed the lowest error (bias – 3.1%, MAD 3.5%), while walking (– 5.9%, MAD 10.8%) and cycling (– 5.3%, MAD 10.1%) also showed relatively close agreement with indirect calorimetry. This is consistent with previous work showing that HR-based estimation performs best during rhythmic exercise, where cardiovascular demand and whole-body metabolic demand are comparatively closely coupled (Achten and Jeukendrup 2003; Schantz et al. 2019). More broadly, a meta-analysis of 60 wearable-validation studies found that EE accuracy varied substantially by activity type and that adding physiological sensing to accelerometry reduced error for most activity types (O'Driscoll et al. 2020).

The larger errors occurred where this coupling becomes weaker. Recovery was overestimated by 15.7%, consistent with HR remaining elevated while oxygen uptake declines toward baseline (Børsheim and Bahr 2003). The bias persisted despite the recovery-specific reference multiplier and relaxation rate, suggesting that a single relaxation rate does not capture between-individual differences in how quickly heart rate and oxygen uptake decouple after maximal exercise, and that further reducing recovery bias would have degraded agreement in other states. Weighted exercise was overestimated by 18.1% and showed relatively high individual error, which is plausible given the localized muscle recruitment and mixed contraction types of resistance exercise. Stair descent produced the largest relative error (+ 35.3%), consistent with the low metabolic cost of eccentric work, although it accounted for very little absolute energy in the protocol. Stair descent represents an extreme eccentric condition in which relatively high muscular force can be generated at low metabolic cost (Douglas et al. 2017; Peñailillo et al. 2017; Renwick et al. 2024).

Estimation without demographic inputs

The third question was whether PHM could estimate light sedentary and recovery EE without age-, sex-, or height-based population equations. Resting metabolic rate (RMR) was not measured; however, light sedentary EE can be regarded as a close, though not equivalent, proxy. PHM estimated light sedentary EE with little systematic error (bias + 0.6%, MAD 9.2%) without explicit demographic predictors. The low sedentary bias partly reflects the tuned anchor; with the framework anchor, Work-phase EE would have been approximately 3% lower. Individual precision was not better than that of conventional RMR equations. The participant-level SD of 11.2% corresponds to 95% limits of roughly $\pm 22\%$, comparable to the limits of about $\pm 19\%$ reported by Kfir et al. (2023) for a newly developed body-composition-based model, across common equations whose mean biases ranged from approximately – 2% to – 17%. Only 10 of 20

participants were within $\pm 10\%$ of reference, whereas Westerterp (2023) reported that the best-performing equation predicted REE within $\pm 10\%$ in just over 70% of adults. PHM thus retains body mass but replaces explicit age, sex, and height terms with individualized physiological quantities including $HR_{\max}$, vVT_2 , and E , with a similar level of individual precision. Because the resting and activity-factor relationships were calibrated in this cohort, a low mean sedentary bias is expected in-sample and is not by itself evidence that demographic terms are unnecessary; whether demographic-free estimation holds with fixed coefficients is the main question for external validation.

PHM without demographic inputs estimated EE during 20 min of passive recovery with a mean bias of + 15.7%. Jung et al. (2021) predicted post-exercise EPOC from fat-free mass and accumulated heart rate, with standard errors corresponding to approximately 18–22% of mean measured EPOC across exercise protocols. Moon et al. (2023) used machine learning and participant characteristics to estimate combined exercise-plus-EPOC EE, achieving correlations with indirect calorimetry of up to 0.976, although recovery accuracy was not reported separately. Although these error metrics are not directly comparable, the present demographic-free recovery result is of a similar order of magnitude to previous EPOC prediction error, but it is a systematic bias rather than random error. Overall, the answer to the third research question is a qualified yes in this development cohort: PHM estimated light sedentary EE with low mean bias but population-equation-level individual error, and recovery EE with moderate systematic overestimation, without using age-, sex-, or height-based population equations. Because both states informed the calibration, this demonstrates feasibility rather than validated demographic-free estimation.

Activity-state detection and wearable implementation

The results from research question 2 also suggest a straightforward route for improving wearable implementation. PHM currently uses broad state-specific mappings for work/other activity, walking, exercise, and recovery. Walking has its own multiplier, but running, cycling, resistance exercise, and stair locomotion share the broader exercise mapping. The different errors observed among those modes suggest that this classification may be too coarse. Accelerometry can distinguish common activity modes such as walking, sitting, standing, running and cycling (Sadeghi Janbahan and Espin-Garcia 2026). A wearable could thus identify activity modes and allow PHM to select a corresponding HR–metabolic mapping to improve EE prediction. This follows the logic of branched HR–accelerometry models (Brage et al. 2004, 2015). This also means that free-living accuracy may depend on activity profile rather than demographic profile. A largely sedentary individual whose activity consists mainly of walking may be easier to model than a frequent cyclist or gym user spending substantial time in modes with less stable HR–EE relationships. That is a testable prediction, not a conclusion from the present study, and should be examined during external free-living validation.

Limitations

The main limitation is that empirical coefficients, including the P_c multiplier and the $HR_{\max}$ offset, were fine-tuned in the same cohort used for evaluation, making the reported performance optimistic relative to independent data. External validation with coefficients fixed a priori is therefore essential. Broad activity-state labels were supplied from the annotated protocol, and automatic detection from wearable sensors was not evaluated. The exercise bouts were relatively short and followed by brief recovery periods, limiting evaluation of EE estimation during prolonged exercise and extended post-exercise recovery. Reference EE was computed from gas exchange with the Weir equation and therefore excludes energy from lactate accumulation and phosphocreatine breakdown, which can be substantial in bouts of 30–60 s and near exhaustion (Briand et al. 2025; Ferretti and di Prampero 2026). Mode-level errors for brief resistance bouts partly reflect this limitation of the criterion.

The individual inputs were not independent of the evaluated data: $HR_{\max}$ and vVT_2 came from the CPET that formed part of the criterion session, and E from its recovery. In practice these inputs would be estimated in the field with additional error, and session EE was sensitive to E (approximately + 3.6% per + 0.05 increase in E ; Online Resource 1), so the

present results represent a best case with respect to input accuracy. The sample was small ($n = 20$), healthy, normal-weight, and age-limited, and the single supervised session did not include sleep or environmental variation. Indirect calorimetry is also subject to breath-by-breath variability and is less reliable for very short bouts (Guidetti et al. 2018; Perez-Suarez et al. 2018). Full wearable validation therefore requires independent cohorts, field-estimated $HR_{\max}/\dot{V}T_{2}/E$, automatic activity-state detection, and ultimately free-living comparison with doubly labelled water (Schoeller et al. 1986), consistent with current recommendations for validation of wearable EE estimation (Argent et al. 2022).

Conclusion

In this development cohort, PHM reproduced total session EE closely, while activity-mode accuracy varied systematically and was highest during rhythmic locomotion. Light sedentary EE was estimated with little mean bias without explicit age-, sex-, or height-based population equations, although individual error was similar to that of such equations, whereas recovery and some resistance or eccentric activities remained more challenging. These findings suggest that individualized heart-rate physiology can provide a useful basis for EE estimation, but that activity context matters. External validation with fixed coefficients, field-estimated inputs, and automatic activity classification is now required, followed by free-living evaluation across different activity profiles.

Abbreviations

AF
activity factor
BF
basal factor
BMR
basal metabolic rate
CPET
cardiopulmonary exercise test
 E
endurance parameter
EE
energy expenditure
 HR_c
mode reference heart rate
 $HR_{\max}$
maximal heart rate
MAP
maximal aerobic power
MAS
maximal aerobic speed
 P_c
resting/response heart-rate anchor
 P_i
HR threshold interval
PHM
physiological heart-rate model
RHR

resting heart rate

$v\dot{V}T_2$

speed at the second ventilatory threshold

$\dot{V}O_{2\max}$

maximal oxygen uptake

$\dot{V}T_2$

second ventilatory threshold

Declarations

Data availability

Participant-level data are not publicly deposited. Requests regarding data access may be directed to the corresponding author and will be considered subject to applicable ethics, participant-consent, and data-protection requirements.

Acknowledgments

The authors thank the laboratory staff who assisted with study procedures and all participants who took part in the study.

Funding

The study was part of the FITSILVER project (Eurostars-3 Project E! 2671), a collaboration between Driftline ehf and CSEM SA from 2023 to 2025, and was co-funded by Innosuisse and Iceland's Technology Development Fund.

Author information - Authors and Affiliations

Faculty of Health Promotion, Sport and Leisure Studies, University of Iceland, Reykjavík, Iceland: Agnar Steinarsson, Gréta Jakobsdóttir, Erlingur Jóhannsson.

Driftline – Applied Physiology and Metabolic Analytics, Reykjavik, Iceland: Jón H. Friðriksson.

Swiss Center of Electronics and Microtechnology (CSEM), Neuchâtel, Switzerland: Guillaume Bonnier, Christopher Moufawad El Achkar, and Alia Lemkaddem.

Author contributions

AS, AL, GB, and CMEA conceived the study and prepared the ethics application. GB and AL recruited participants and collected data. AS, JHF, and GB analyzed the data. AL supervised the study. AS prepared the initial and final drafts of the manuscript. EJ and GJ supervised manuscript preparation for publication. All authors read, revised, and approved the final version of the manuscript.

Corresponding author

Correspondence to Erlingur Jóhannsson.

Ethics declarations

Conflict of interest

AS, the PhD candidate supervised by EJ and GJ, is the founder of Driftline, a company developing heart-rate-based fitness and metabolic tracking technologies. AS is the inventor of the physiological heart-rate model described in this study. A PCT patent application covering aspects of this method has been filed. JHF was employed by Driftline at the time of the study. The remaining authors declare no commercial or financial relationships that could be construed as a potential conflict of interest. The study was conducted under academic supervision (EJ and GJ) and approved by the relevant institutional bioethics committee.

Ethical approval

The study was approved by the Swiss Ethics CER-VD (BASEC ID 2024-01496) and conducted in accordance with local legislation, institutional requirements, and the Declaration of Helsinki. All participants provided written informed consent to participate.

References

1. Achten J, Jeukendrup AE (2003) Heart rate monitoring: applications and limitations. *Sports Med* 33:517–538. <https://doi.org/10.2165/00007256-200333070-00004>
2. Argent R, Hetherington-Rauth M, Stang J, Tarp J, Ortega FB, Molina-Garcia P, Schumann M, Bloch W, Cheng S, Grøntved A, Brønd JC, Ekelund U, Sardinha LB, Caulfield B (2022) Recommendations for determining the validity of consumer wearables and smartphones for the estimation of energy expenditure: expert statement and checklist of the INTERLIVE Network. *Sports Med* 52:1817–1832. <https://doi.org/10.1007/s40279-022-01665-4>
3. Binder RK, Wonisch M, Corra U, Cohen-Solal A, Vanhees L, Saner H, Schmid JP (2008) Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *Eur J Cardiovasc Prev Rehabil* 15:726–734. <https://doi.org/10.1097/HJR.0b013e328304fed4>
4. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
5. Briand J, di Prampero PE, Osgnach C, Thibault G, Tremblay J (2025) Quantifying metabolic energy contributions in sprint running: a novel bioenergetic model. *Eur J Appl Physiol* 125:3521–3541. <https://doi.org/10.1007/s00421-025-05831-0>
6. Børsheim E, Bahr R (2003) Effect of exercise intensity, duration and mode on post-exercise oxygen consumption. *Sports Med* 33:1037–1060. <https://doi.org/10.2165/00007256-200333140-00002>
7. Brage S, Brage N, Franks PW, Ekelund U, Wong MY, Andersen LB, Froberg K, Wareham NJ (2004) Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J Appl Physiol* 96:343–351. <https://doi.org/10.1152/jappphysiol.00703.2003>
8. Brage S, Westgate K, Franks PW, Stegle O, Wright A, Ekelund U, Wareham NJ (2015) Estimation of free-living energy expenditure by heart rate and movement sensing: a doubly-labelled water study. *PLoS ONE* 10:e0137206. <https://doi.org/10.1371/journal.pone.0137206>
9. Chen KY, Bassett DR Jr (2005) The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc* 37:S490–S500. <https://doi.org/10.1249/01.mss.0000185571.49104.82>
10. Douglas J, Pearson S, Ross A, McGuigan M (2017) Eccentric exercise: physiological characteristics and acute responses. *Sports Med* 47:663–675. <https://doi.org/10.1007/s40279-016-0624-8>
11. Ferretti G, di Prampero PE (2026) A reassessment of the energetic significance of blood lactate accumulation during exercise. *Eur J Appl Physiol* 126:1279–1286. <https://doi.org/10.1007/s00421-026-06134-8>
12. Guidetti L, Meucci M, Bolletta F, Emerenziani GP, Gallotta MC, Baldari C (2018) Validity, reliability and minimum detectable change of the COSMED K5 portable gas exchange system in breath-by-breath mode. *PLoS ONE*

13:e0209925. <https://doi.org/10.1371/journal.pone.0209925>

13. Jung W-S, Park H-Y, Kim S-W, Kim J, Hwang H, Lim K (2021) Estimating excess post-exercise oxygen consumption using multiple linear regression in healthy Korean adults: a pilot study. *Phys Act Nutr* 25(1):35–41. <https://doi.org/10.20463/pan.2021.0006>
14. Keir DA, Pogliaghi S, Inglis EC, Murias JM, Iannetta D (2024) The respiratory compensation point: mechanisms and relation to the maximal metabolic steady state. *Sports Med* 54:2993–3003. <https://doi.org/10.1007/s40279-024-02084-3>
15. Keytel LR, Goedecke JH, Noakes TD, Hiiloskorpi H, Laukkanen R, van der Merwe L, Lambert EV (2005) Prediction of energy expenditure from heart rate monitoring during submaximal exercise. *J Sports Sci* 23:289–297. <https://doi.org/10.1080/02640410470001730089>
16. Kfir A, Lahav Y, Gepner Y (2023) Cross-validation of a new general population resting metabolic rate prediction equation based on body composition. *Nutrients* 15:805. <https://doi.org/10.3390/nu15040805>
17. Lin LI (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255–268. <https://doi.org/10.2307/2532051>
18. Lyden K, Kozey SL, Staudenmeyer JW, Freedson PS (2011) A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. *Eur J Appl Physiol* 111:187–201. <https://doi.org/10.1007/s00421-010-1639-8>
19. Moon J, Oh M, Kim S, Lee K, Lee J, Song Y, Jeon JY (2023) Intelligent estimation of exercise induced energy expenditure including excess post-exercise oxygen consumption (EPOC) with different exercise intensity. *Sensors* 23:9235. <https://doi.org/10.3390/s23229235>
20. O'Driscoll R et al (2020) How well do activity monitors estimate energy expenditure? A systematic review and meta-analysis of the validity of current technologies. *Br J Sports Med* 54:332–340. <https://doi.org/10.1136/bjsports-2018-099643>
21. Peñailillo L, Blazevich AJ, Nosaka K (2017) Factors contributing to lower metabolic demand of eccentric compared with concentric cycling. *J Appl Physiol* 123:884–893. <https://doi.org/10.1152/jappphysiol.00536.2016>
22. Perez-Suarez I, Martin-Rincon M, Gonzalez-Henriquez JJ, Fezzardi C, Perez-Regalado S, Galvan-Alvarez V, Juan-Habib JW, Morales-Alamo D, Calbet JAL (2018) Accuracy and precision of the COSMED K5 portable analyser. *Front Physiol* 9:1764. <https://doi.org/10.3389/fphys.2018.01764>
23. Rennie KL, Hennings SJ, Mitchell J, Wareham NJ (2001) Estimating energy expenditure by heart-rate monitoring without individual calibration. *Med Sci Sports Exerc* 33:939–945. <https://doi.org/10.1097/00005768-200106000-00013>
24. Renwick NC, Egginton S, Ferguson C (2024) Reduced cardiovascular and metabolic responses during eccentric stepping exercise: A pilot study. *Physiol Rep* 12:e70080. <https://doi.org/10.14814/phy2.70080>
25. Sadeghi Janbahan K, Espin-Garcia O (2026) Methods for classifying physical activities using accelerometer data: a scoping review. *npj Digit Med* 9:534. <https://doi.org/10.1038/s41746-026-02694-3>
26. Sardinha LB, Júdice PB (2017) Usefulness of motion sensors to estimate energy expenditure in children and adults: a narrative review of studies using doubly labelled water. *Eur J Clin Nutr* 71:331–339. <https://doi.org/10.1038/ejcn.2017.2>
27. Schantz P, Salier Eriksson J, Rosdahl H (2019) The heart rate method for estimating oxygen uptake: analyses of reproducibility using a range of heart rates from commuter walking. *Eur J Appl Physiol* 119:2655–2671. <https://doi.org/10.1007/s00421-019-04236-0>
28. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, Jéquier E (1986) Energy expenditure by doubly labeled water: validation in humans and proposed calculation. *Am J Physiol* 250:R823–R830.

29. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, Wheeler MT, Ashley EA (2017) Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *J Pers Med* 7:3. <https://doi.org/10.3390/jpm7020003>
30. Spurr GB, Prentice AM, Murgatroyd PR, Goldberg GR, Reina JC, Christman NT (1988) Energy expenditure from minute-by-minute heart-rate recording: comparison with indirect calorimetry. *Am J Clin Nutr* 48:552–559. <https://doi.org/10.1093/ajcn/48.3.552>
31. Swain DP, Leutholtz BC (1997) Heart rate reserve is equivalent to %VO₂ reserve, not to %VO₂max. *Med Sci Sports Exerc* 29:410–414. <https://doi.org/10.1097/00005768-199703000-00018>
32. The jamovi project (2025) jamovi (Version 2.6.44) [Computer software]. <https://www.jamovi.org>
33. de Weir JB V (1949) New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol* 109:1–9. <https://doi.org/10.1113/jphysiol.1949.sp004363>
34. Westerterp KR (2023) Predicting resting energy expenditure: a critical appraisal. *Eur J Clin Nutr* 77:953–958. <https://doi.org/10.1038/s41430-023-01299-3>
35. Wicks JR, Oldridge NB, Nielsen LK, Vickers CE (2011) HR index-a simple method for the prediction of oxygen uptake. *Med Sci Sports Exerc* 43:2005–2012. <https://doi.org/10.1249/MSS.0b013e318217276e>

Figures

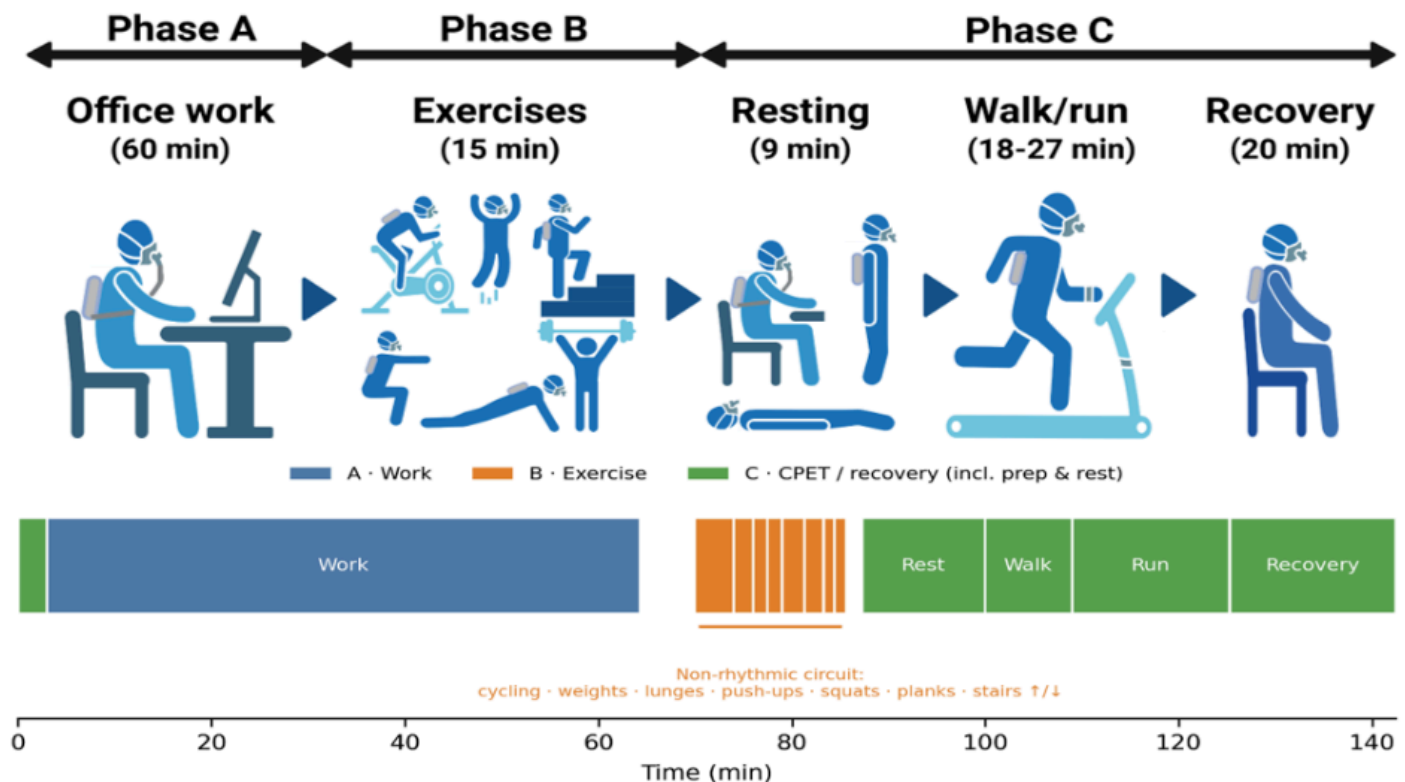


Figure 1

Experimental protocol. Representative images and timeline of the single-session protocol, comprising three phases: Phase A (Work, ~1 h of seated/standing office work activity), Phase B (Exercise, a non-rhythmic circuit of cycling, bicep curls, lunges, push-ups, squats, planks, and stair ascent/descent), and Phase C (CPET/recovery, comprising preparation,

seated rest, treadmill walking, treadmill running to volitional exhaustion, and seated recovery). Blocks are colored by phase and drawn to scale in time. Created in BioRender. Steinarsson, A. (2026) <https://BioRender.com/saqamnx>

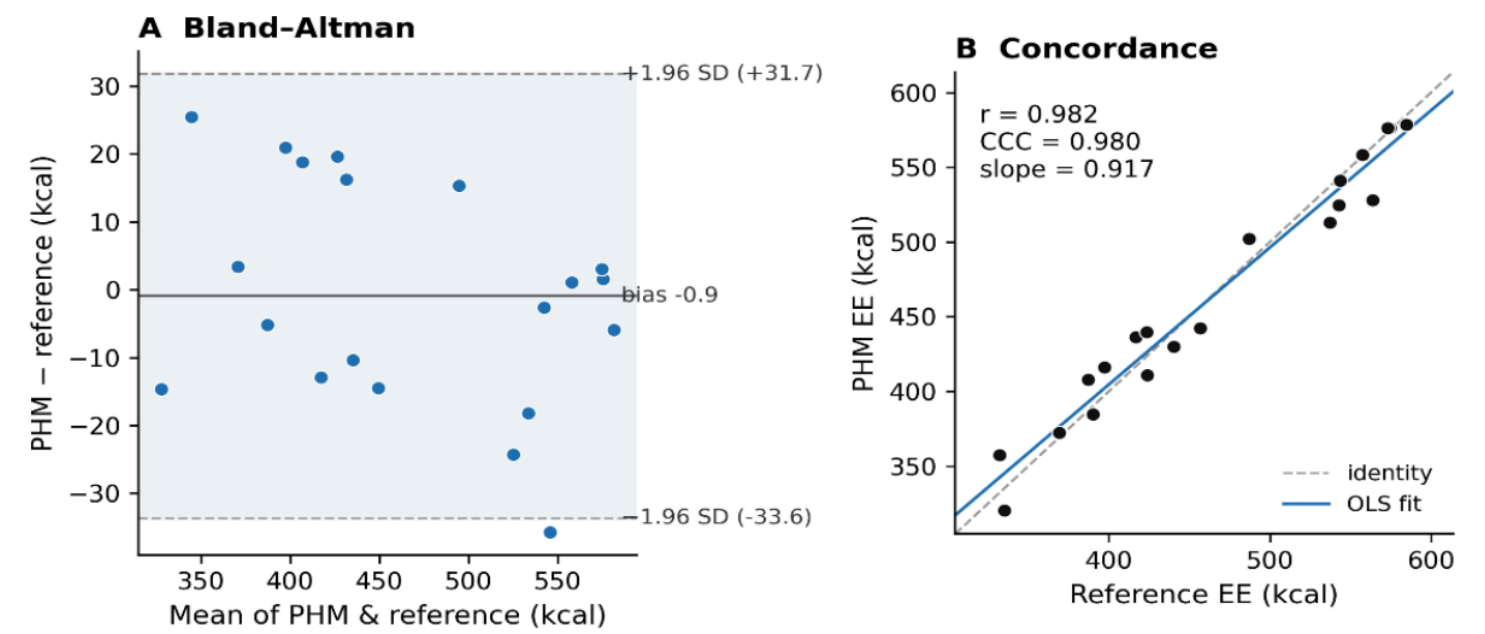


Figure 2

Whole-session agreement between PHM-estimated and reference energy expenditure ($n = 20$). (A) Bland–Altman plot of per-participant session totals; the solid line is the mean bias (-0.9 kcal) and dashed lines the 95% limits of agreement (-33.6 to $+31.7$ kcal). (B) PHM versus reference session totals with the line of identity (dashed) and the ordinary-least-squares fit (solid); Pearson $r = 0.98$, Lin's concordance correlation coefficient = 0.98 , slope = 0.92 .

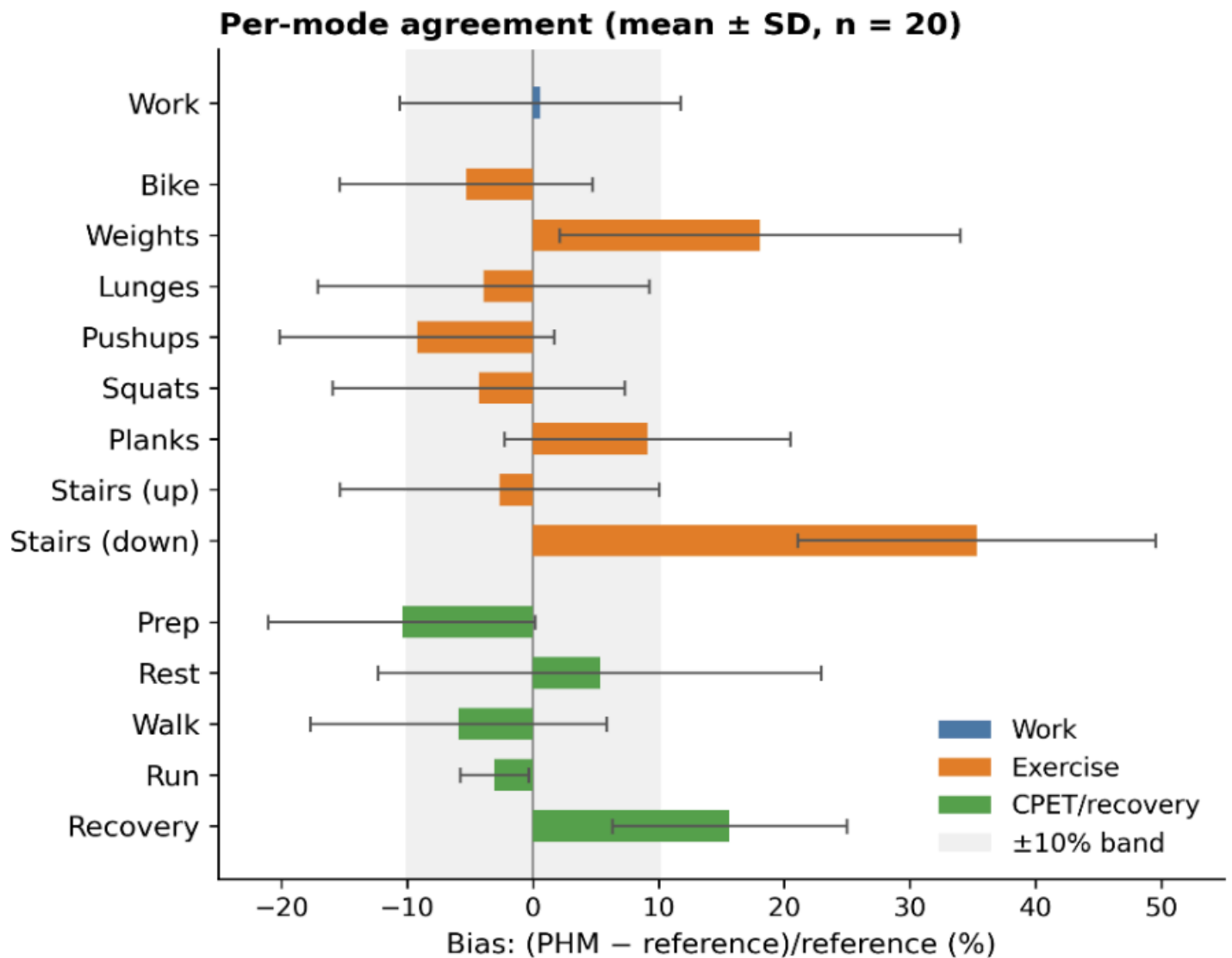


Figure 3

Agreement between PHM-estimated and reference energy expenditure by activity mode. Mean bias (PHM – ref)/ref, expressed as a percentage, for each activity mode; error bars are $\pm$ SD across participants ($n = 20$). The shaded band marks $\pm 10\%$.

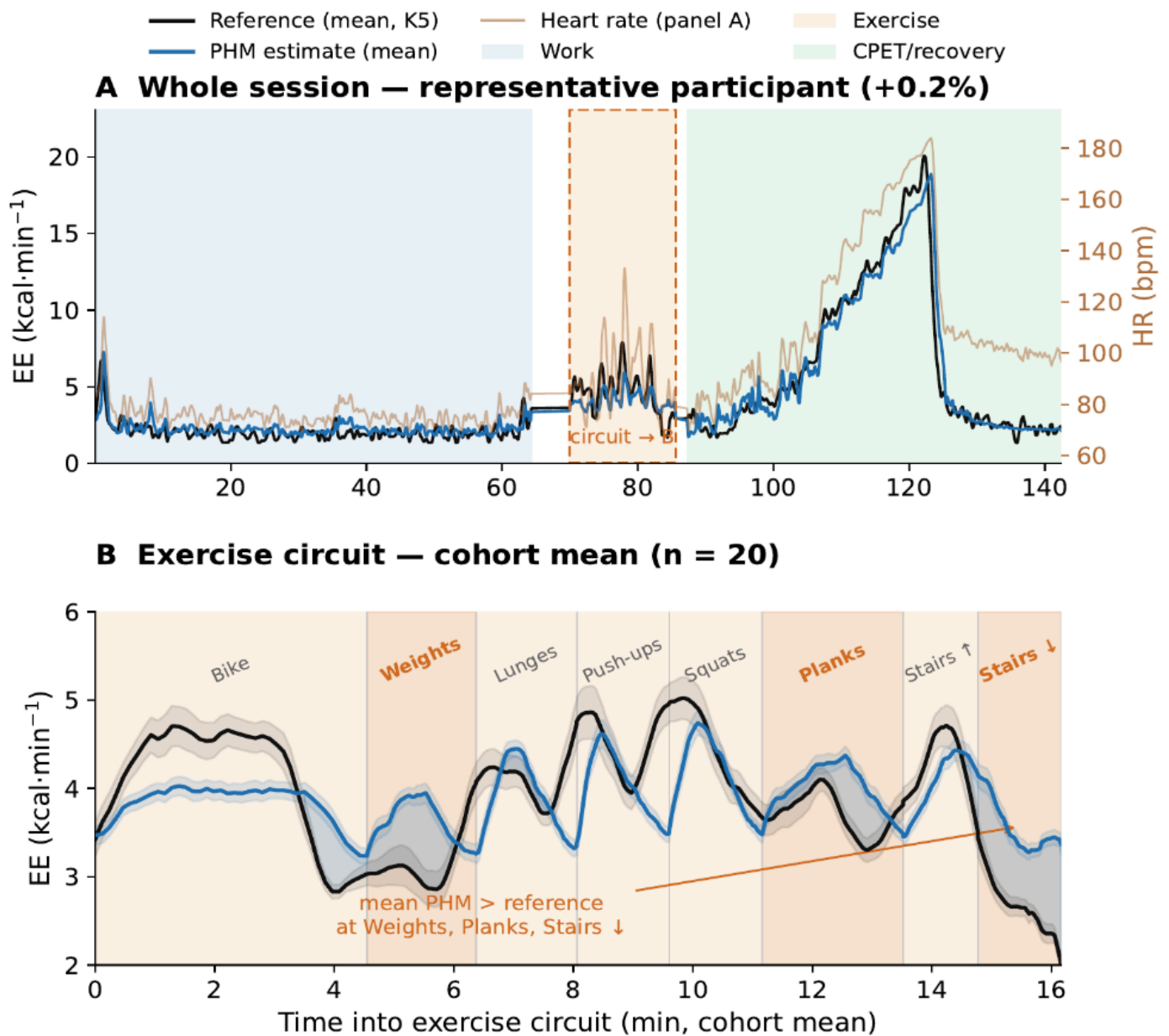


Figure 4

Second-by-second energy expenditure from the physiological heart-rate model (PHM, blue) and indirect calorimetry (reference; Cosmed K5, black). (A) Whole session for a representative participant (deviation +0.2%); heart rate (right axis) and protocol-phase shading shown for context; dashed box marks the exercise circuit in (B). (B) The exercise circuit averaged across all participants (cohort mean, $n = 20$), aligned by activity mode and plotted against mean elapsed circuit time; shaded bands are ± 1 SE.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SteinarssonEEpaperESM1.pdf](#)