Automatic Heart Rate Normalization for Accurate Energy Expenditure Estimation

An Analysis of Activities of Daily Living and Heart Rate Features

M. Altini1,2; J. Penders1; R. Vullers1; O. Amft1,3
1Holst Centre/imec The Netherlands, Eindhoven, The Netherlands; 2Eindhoven University of Technology, Eindhoven, The Netherlands; 3ACTLab, University of Passau, Passau, Germany

Keywords
Energy expenditure, heart rate, wearable sensors, personalization

Summary
Introduction: This article is part of the Focus Theme of Methods of Information in Medicine on “Pervasive Intelligent Technologies for Health”.

Background: Energy Expenditure (EE) estimation algorithms using Heart Rate (HR) or a combination of accelerometer and HR data suffer from large error due to inter-person differences in the relation between HR and EE. We recently introduced a methodology to reduce inter-person differences by predicting a HR normalization parameter during low intensity Activities of Daily Living (ADLs). By using the HR normalization, EE estimation performance was improved, but conditions for performing the normalization automatically in daily life need further analysis. Sedentary lifestyle of many people in western societies urge for an in-depth analysis of the specific ADLs and HR features used to perform HR normalization, and their effects on EE estimation accuracy in participants with varying Physical Activity Levels (PALs).

Objectives: To determine 1) which low intensity ADLs and HR features are necessary to accurately determine HR normalization parameters, 2) whether HR variability (HRV) during ADLs can improve accuracy of the estimation of HR normalization parameters, 3) whether HR normalization parameter estimation from different ADLs and HR features is affected by the participants’ PAL, and 4) what is the impact of different ADLs and HR features used to predict HR normalization parameters on EE estimation accuracy.

Methods: We collected reference EE from indirect calorimetry, accelerometer and HR data using one single sensor placed on the chest from 36 participants while performing a wide set of activities. We derived HR normalization parameters from individual ADLs (lying, sedentary, walking at various speeds), as well as combinations of sedentary and walking activities. HR normalization parameters were used to normalized HR and estimate EE.

Results: From our analysis we derive that 1) HR normalization using resting activities alone does not reduce EE estimation error in participants with different reported PALs. 2) HRV features did not show any significant improvement in RMSE. 3) HR normalization parameter estimation was found to be biased in participants with different PALs when sedentary-only data was used for the estimation. 4) EE estimation error was not reduced when normalization was carried out using sedentary activities only. However, using data from walking at low speeds improved the results significantly (30–36%).

Conclusion: HR normalization parameters able to reduce EE estimation error can be accurately estimated from low intensity ADLs, such as sedentary activities and walking at low speeds (3–4 km/h), regardless of reported PALs. However, sedentary activities alone, even when HRV features are used, are insufficient to estimate HR normalization parameters accurately.

1. Introduction
1.1 Scientific Background
Physical Activity (PA) and exercise capacity are among the most important determinants of health, both in healthy subjects as well as disease populations [1]. The reduced Physical Activity Level (PAL) in most western countries is causing new epidemics such as obesity and diabetes to spread [2]. Ubiquitous technologies, such as accelerometers and Heart Rate (HR) monitors [3], started providing unprecedented insights into links of PA and health. Early epidemiological research focused on developing single models or branched equations combining accelerometer and HR data to predict EE [4–6]. These ap-
proaches are motivated by the relations between body movement and EE as well as between oxygen intake, HR and EE. The limitation of these methods include that a single accelerometer worn close to the body center of mass cannot detect low and upper body motion [7], the reduced relevance of HR during sedentary behavior and the need for individual calibration [6]. By introducing activity-specific models, consisting of a two-step process, where first an activity is recognized, and then an EE estimation model is applied, researchers were able to tackle some of these limitations [3, 5]. The relation between EE and acceleration as well as HR is peculiar of a specific context (e.g. activity), thus activity-specific models are able to capture this relation beyond what single regression models or branched models can do [7–9]. Even though algorithms including HR consistently provided improvements compared to accelerometers alone [4, 6, 9], the main limitation of HR – which is the need for individual calibration – requires a different solution. Decomposing the EE estimation process into activity-specific sub-problems is not sufficient to take into account the different relation between HR and EE in different individuals.

During moderate to vigorous PA, differences in HR between persons performing the same activity are mainly due to cardiorespiratory fitness (CRF). However, differences in CRF level do not cause different metabolic responses [10]. Nevertheless, CRF-related variance was tackled only by means of individual calibration [6] and/or by performing intense activities such as running [11]. For many practical applications personal calibration is not feasible since it would require every user to perform a suitable fitness test. We recently introduced a methodology to automatically normalize HR by estimating a normalization parameter that describes HR at a certain workload, using low intensity ADLs [13] (see Fig. 1). The methodology is based on the tight relation between CRF and the HR at a certain workload, which is the basis of sub-maximal CRF tests [14].

### 1.2 Rationale for the Study

Practical conditions for performing the normalization automatically in daily life need further analysis. The sedentary lifestyle of many people in western societies [15] urge for an in-depth analysis of the specific ADLs required to predict HR normalization parameters, and their effects on EE estimation accuracy in persons with varying PALs. Additionally, HR variability (HRV) features from sedentary activities as well as moderate to intense ones have been shown to be linked to CRF level and PALs in past research [16, 17]. Even though this link is unclear, and results are often in disagreement [18–20], given the close relation between CRF and HR normalization parameters it is of interest to analyze if HRV features can predict HR normalization parameters and reduce EE estimation error.

### 1.3 Objectives of the Study

This is the first analysis of how low intensity ADLs and HR features can be used to estimate HR normalization parameters, and their effects on EE estimation accuracy. Our objectives are: 1) To determine which ADLs and HR features are necessary to accurately determine HR normalization parameters, 2) To determine whether HRV during ADLs can improve accuracy of the estimation of HR normalization parameters, 3) To determine whether HR normalization parameters estimation from different ADLs and HR features is affected by the participants’ PAL and 4) To determine what is the impact of different ADLs and HR features used to predict HR normalization parameters on EE estimation accuracy.

### 2. Methods

#### 2.1 Participants

Participants were 36 (27 male, 9 female) self-reported healthy Holst Centre employees from diverse ethnic background. Mean age was 31.2 ± 5.7 years, mean weight was 73.3 ± 11.2 kg, mean height was 176.6 ± 9.1 cm and mean BMI was 23.4 ± 2.4 kg/m². Imec’s IRB approved the study, and each participant signed an informed consent form.

#### 2.2 Study Design

Participants reported at the lab after refraining from drinking (except for water), eating and smoking in the two hours before the experiment. The protocol consisted of common ADLs in industrialized countries [22], as well as intense activities. Activities were grouped into six clusters to be used for activity classification. The six clusters were lying (lying down), sedentary (sitting, standing, desk work, reading, writing, PC work, watching TV), dynamic (stacking groceries, washing dishes, cooking, folding clothes, sweeping, vacuuming), walking (treadmill flat at 3, 4, 5, 6 km/h, inclined 3–5%, 3–5 km/h), biking (low medium and high resistance level at 60 and 80 rpms), running (treadmill 7, 8, 9, and 10 km/h). Activities lasted for a period of at least 4 minutes, with the exception of running (1 to 4 minutes).

#### 2.3 Outcome Measures

All analyses were performed independent of the participant (leave one subject out validation). Accuracy of the HR normalization parameter estimation was evaluated using: 1) Pearson’s correlation between each HR feature and the HR normalization parameter, to determine the predictive power of each single feature in each ADLs, 2) the error derived from the difference between estimated and measured normalization parameters, to determine possible bias and precision of the estimate. As the measured normalization parameter we used the actual HR while running on a treadmill. 3) The Root Mean Square Error (RMSE) between estimated and measured normalization parameters, to determine the accuracy of the estimate. Additionally, participants were split in active (ACT) and inactive (INA) groups, based on reported PALs in order to determine possible PAL-induced bias in the estimation procedure. The performance measure used for EE was the RMSE, averaged within an activity and between participants. A one-way repeated-measures within-subjects ANOVA with five levels was used to compare RMSE between EE models. The Tukey test was used as a post hoc test to perform pairwise comparisons and identify significant differ-
ences. In addition, unpaired t-tests were used to compare INA and ACT groups. Significance was assessed at $\alpha < 0.05$ for all analyses.

2.4 Methods for Data Acquisition and Measurement

2.4.1 ECG Necklace

The ECG Necklace [23] is a low power wireless ECG platform which was configured to acquire one lead ECG data at 256 Hz, and accelerometer data from a three-axial accelerometer at 32 Hz (▶Figure 2). The sensor was placed on the chest with an elastic belt. Two gel electrodes were placed on the participant’s chest, in the lead II configuration.

2.4.2 Indirect Calorimeter

Breath-by-breath data were collected using the Cosmed K4b2 indirect calorimeter. The Cosmed K4b2 weights 1.5 kg including battery and showed to be a reliable measure of EE [24].

2.5 Methods for Data Analysis

Accelerometer and HR features were used to derive activity recognition models, walking speed, HR normalization parameter estimation models and EE estimation linear models (▶Figure 1). To estimate walking speed, we deployed multiple regression models using accelerometer-only features as predictors according to [12, 13]. Details on the accelerometer features and on the implementation of the models have been widely covered elsewhere [9, 13]. Here, we will focus on the HR features and ADLs used for the estimation of the HR normalization parameter.

2.5.1 HR Features

HR features were extracted from R-R intervals, computed over 2 minutes windows to ensure sufficient frequency resolution in the Low Frequency band [25]. Time domain features included mean HR (meanHR), standard deviation of beat-to-beat intervals (SDNN), square root of the mean squared difference of successive R-Rs.
Normalized Heart Rate to Estimate Energy Expenditure (rMSSD) and number of pairs of successive R-Rs that differ by more than 50 ms (pNN50). Frequency domain features included low (LF, 0.04–0.15 Hz) and high frequency power (HF, 0.15–0.40 Hz).

2.5.2 Automatic HR Normalization Parameter Estimation from ADLs

Multiple linear regression models were built to analyze individual ADLs that can be recognized with high recognition rates (e.g. lying 100%, sedentary 91% and walking 98%, together with walking speed – RMSE 0.28 ± 0.09 km/h [13]), as well as combinations of such ADLs. For each ADL we built a multiple linear regression model using as predictors HR and/or HRV features during such ADL, and as dependent variable the HR normalization parameter. As reference HR normalization parameter we selected running at 9 km/h, which was of sufficient intensity to provide precise HR normalization. No performance improvement in EE estimation accuracy was found in our dataset when using more intense workloads to normalize HR. The activities and combinations of activities selected as ADL to predict the HR normalization parameter were the following:

- **Lying**: lying down resting
- **Sed**: sedentary activities
- **Walk 3–4–5–6**: walking at 3–4–5 or 6 km/h
- **Comb A**: Lying, Sed, Walk3 and Walk4
- **Comb B**: Lying, Sed, Walk3, Walk4, Walk5 and Walk6

To analyze the impact of HRV features, two multiple regression models were built for each activity and combination, one including HR only, and one including HR and HRV features.

2.5.3 EE Estimation

EE was estimated by first classifying the activity performed using accelerometer features and then applying an activity-specific EE linear regression model. The activity-specific EE linear models use anthropometric characteristics, accelerometer and HR features. Thus, we developed six multiple linear regression models, one for each cluster. Normalized HRs (i.e. HR divided by the estimated HR normalization parameter) obtained from different sets of ADLs (see 2.5.2) were used as predictors in the multiple regression models for moderate to vigorous clusters (dynamic, walking, running and biking).

![Figure 3](image-url)
3. Results

3.1 Automatic HR Normalization Parameter Estimation from ADLs

Mean HR showed significant correlation with the HR normalization parameter during all ADLs (lying 0.50, sed 0.50, walk3 km/h 0.86, walk4 km/h 0.86, walk5 km/h 0.88 and walk6 km/h 0.90, $p < \alpha$). No HRV feature was found significantly correlated to the HR normalization parameter, in any ADL analyzed ($p > \alpha$ for each HRV feature in each ADL). Additionally, no HRV feature was able to discriminate between participants groups divided by PALs (INA vs ACT), in any activity except for low speed walking ($p < \alpha$ for SDNN, pNN50, LF and HF during walk3). Mean HR could discriminate between INA and ACT in all activities ($p < \alpha$).

Figure 3a–c shows the density plot of the difference between estimated and measured HR normalization parameters. The spread of the distribution reduced by 47% from lying to walk6. Figure 3d–i shows the difference distribution for models where HR or HR + HRV features were predictors, for single ADLs. No significant difference was found when including HRV features in any activity. RMSE was 17.6 bpm for lying data, 17.6 for sed, 10.5 for walk3, 10.3 for walk4, 10.4 for walk5, 9.4 for walk6, 11.8 for CombA and 9.0 for CombB. No differences in RMSE were found when HR and HRV features were combined ($p > \alpha$ for all activities).

Figure 3j shows the HR normalization parameter estimation error when different ADLs are used as predictors, divided per PAL of the participants. When only resting data is used (e.g. lying), the HR normalization parameter is overestimated for ACT participants, while it is underestimated for INA ones. No difference in the estimation accuracy between ACT and INA participants was found when walking data was included in the models as well (CombA and CombB), with higher walking speeds (CombB) showing higher precision (spread further reduced by 24%).

3.2 EE Estimation

Figure 4 shows the results of the HR normalization on EE estimation. The results of three different normalizations (from lying data only and using combined lying and walking speed data, CombA and CombB), was compared to the cases of no normalization (No Norm) and normalization using the measured HR normalization parameter (Opt Norm) is also shown for comparison. The first column of each subplot shows performance of state of the art activity-specific EE models combining accelerometer and heart rate features, but without HR normalization.

![Figure 4](image-url) Algorithm performance in terms of RMSE of EE estimations during different moderate to vigorous activity clusters (a to d show dynamic activities, e.g. household, and walking, biking and running activities) where HR is not normalized (No Norm), normalized using lying data only, and normalized using ADLs included in CombA and CombB. Normalization performed using the measured HR normalization parameter (Opt Norm) is also shown for comparison. The first column of each subplot shows performance of state of the art activity-specific EE models combining accelerometer and heart rate features, but without HR normalization.
the measured HR normalization parameter (Opt Norm). State of the art activity-specific EE models combining accelerometer and heart rate features were used for all analysis. The only model not including HR normalization is No Norm. No significant error reduction was shown when the HR normalization parameter estimated using lying data only was used (6 to 17%, \( p > \alpha \)). Error reduction when walking data was included was significant for walking activities (36 –37%, \( p < \alpha \), biking activities (30 –38%, \( p < \alpha \)) and running activities (31 –40%, \( p < \alpha \)), but not for dynamic activities (14–15%, \( p > \alpha \)). CombA and CombB could reduce RMSE at the same extent the optimal HR normalization could (difference between CombA, CombB and Opt Norm was not statistically significant, \( p > \alpha \)).

4. Discussion

4.1 Answers to Study Questions

We report the main findings of our analysis, in relation to the four objectives of this study. 1) To determine which ADLs and HR features are necessary to accurately determine HR normalization parameters: from our analysis we derive that resting activities alone are not sufficient to estimate HR normalization parameter, even if there is positive correlation between HR at rest and the HR normalization parameter. Thus, resting activities alone are unable to reduce EE estimation error in participants with different reported PALs. However, results obtained using data at rest and walking at low speeds (e.g. \( \leq 4 \text{ km/h} \)), showed results comparable to the ones obtained when including data while walking at higher speeds. Hence, ADL and HR features support estimating the HR normalization parameter in typical mixed lifestyle. 2) To determine whether HRV during ADLs can improve accuracy of the estimation of HR normalization parameters: from our analysis HRV features were unable to provide additional information and therefore improve the estimate accuracy of the HR normalization parameter (Figure 3d –i). We attribute this finding to a weaker inter-personal relation between HRV and CRF. 3) Whether HR normalization parameter estimation from different ADLs and HR features is affected by the participants’ PAL: our analysis showed that the normalization procedure works equally well in participants with different PALs, provided that walking data is included in the HR normalization parameter multiple linear regression models (Figure 3j). Estimating precision is improved when data while walking at higher speeds is included in the HR normalization parameter multiple linear regression models (see Sec. 2.5.2 and Figure 3a–c). 4) To determine what is the impact of different ADLs and HR features used to predict HR normalization parameters on EE estimation accuracy: our analysis showed that EE estimation accuracy when the HR normalization parameter is estimated from ADLs including walking (CombA and CombB) reaches the same accuracy of the optimal normalization that could be performed measuring the HR normalization parameter during a treadmill test (Figure 4).

4.2 Strength and Weaknesses of the Study

To the best of our knowledge, this is the first time that HR and HRV features are investigated during ADLs as predictors of a HR normalization parameter, together with the impact of such normalization procedure on EE estimation accuracy and participants with different PALs. Using the proposed personalization approach, it is possible to significantly reduce EE estimation error by automatically normalizing HR using low intensity ADLs, such as sedentary activities and walking at low speeds. However, we recognize limitations in our study: Even though we developed algorithms able to derive the HR normalization parameter automatically during ADLs, we tested it using laboratory recordings only. We consider that evaluation with lab data is a necessary first step. In particular, the approach allowed us to establish the accuracy of EE estimation models derived with ADLs and HR features. Further investigations should explore the relation between specific contexts and physiological parameters beyond linear models. The analysis should also be extended to a wider population consisting of participants with varying cardiorespiratory fitness level.

4.3 Results in Relation to Other Studies

Previous work by our group [13] as well as others [8, 11] showed that normalizing the HR using a normalization parameter representative of CRF, such as the HR at a certain workload, can significantly reduce inter-person differences and consequently improve EE estimation accuracy. However, to determine the HR normalization parameter for an individual, required personal calibration (e.g. performing a treadmill test), which is not practical. Moreover, the calibration would need to be repeated frequently. In this study we investigated the possibility to determine the HR normalization parameter from different combinations of ADLs, including rest only activities (e.g. lying or sedentary). Additionally, we analyzed HRV features during ADLs, in the context of EE estimation.

Given the tight relation between CRF and the HR normalization parameter, which is the basis of sub-maximal CRF tests [14], it is of interest to review previous research on the relation between HRV and CRF. Many studies investigated the relation between HR and CRF during cross-sectional studies [18 –20], as well as interventions [26, 27], and showed reductions in HR due to higher CRF levels, but no changes in HRV. Our results are in agreement with those, where HRV features could explain very little of the differences in fitness level, and mean HR was the best predictor of such differences. Since differences in HR and HRV features at rest are mainly driven by age, while feature differences during exercise are mainly driven by fitness [28], we investigated HRV during low intensity ADLs as well. However, we could not find a relation between HRV features while walking and the HR normalization parameter. Other authors did report a significant increase in HRV features and CRF following a physical activity intervention [29], however it is not clear if HRV features could be used as predictors of CRF.
5. Conclusions

We analyzed the impact of HR and HRV features in different ADLs as predictors of a HR normalization parameter necessary in order to reduce inter-person differences in HR and improve EE estimation accuracy. Using HR and HRV features during ADLs as predictors, we aimed at providing a normalization procedure able to automatically normalize HR without requiring any specific test. Overall, we conclude that an accurate personalized EE estimation is feasible, even when only data at rest and from walking at low speeds is available, as frequently occurring in today’s lifestyle.

References