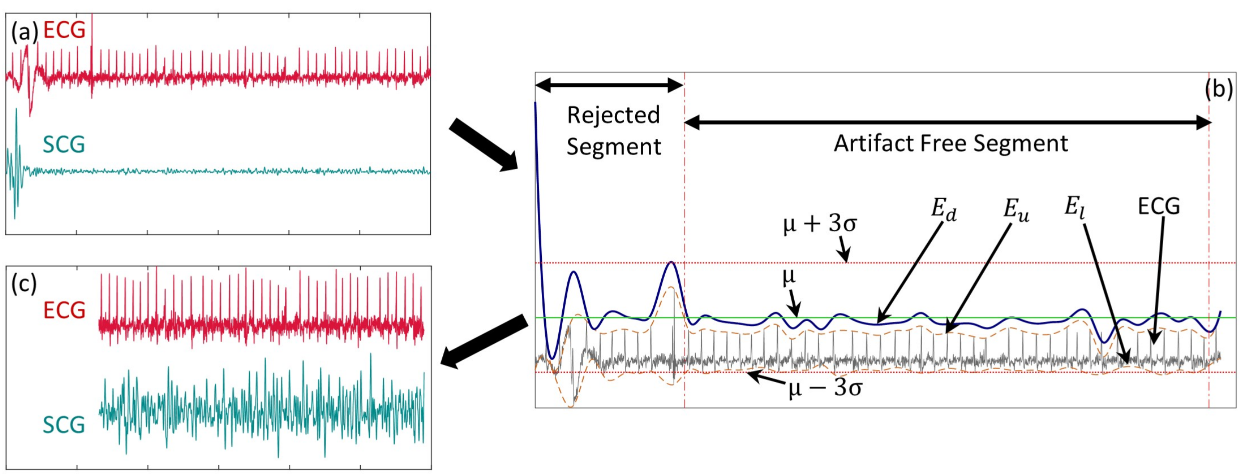
Wearable Patch Based Estimation of Oxygen Uptake and Assessment of Clinical Status during Cardiopulmonary Exercise Testing in Patients with Heart Failure

Md Mobashir Hasan Shandhi, Sinan Hersek, Joanna Fan, Erica Sander, Teresa De Marco, Alex Heller, Mozziyar Etemadi, Liviu Klein and Omer T. Inan

Supplemental Materials

Artifact Free ECG Frame Detection: Motion artifacts in the ECG signal can pose a major challenge to subsequent signal processing and regression steps. Accordingly, an algorithm was implemented to automatically detect, and remove, artifact-corrupted segments of the signal. Specifically, the ECG for each 30-second-long frame was passed through an artifact detection function to choose the window frame (i.e. length) of signal that is artifact free (Figure S1). First, the upper (E_u) and lower (E_l) envelope of the data is detected and a difference E_d (= $E_u - E_l$) is computed. Then, the mean (μ) and standard deviation (σ) of E_d throughout the recording is calculated. An artifact is defined as the signal segment when the E_d of that specific portion is greater or less than 3σ from the μ. The longest artifact-free segment of the signal was chosen and the time stamp for this portion was used to segment all the wearable signals to obtain the signals where high quality ECG was present. Each frame was then visually inspected to verify that the algorithm was successfully removing the artifact-affected ECG portion from the signal.



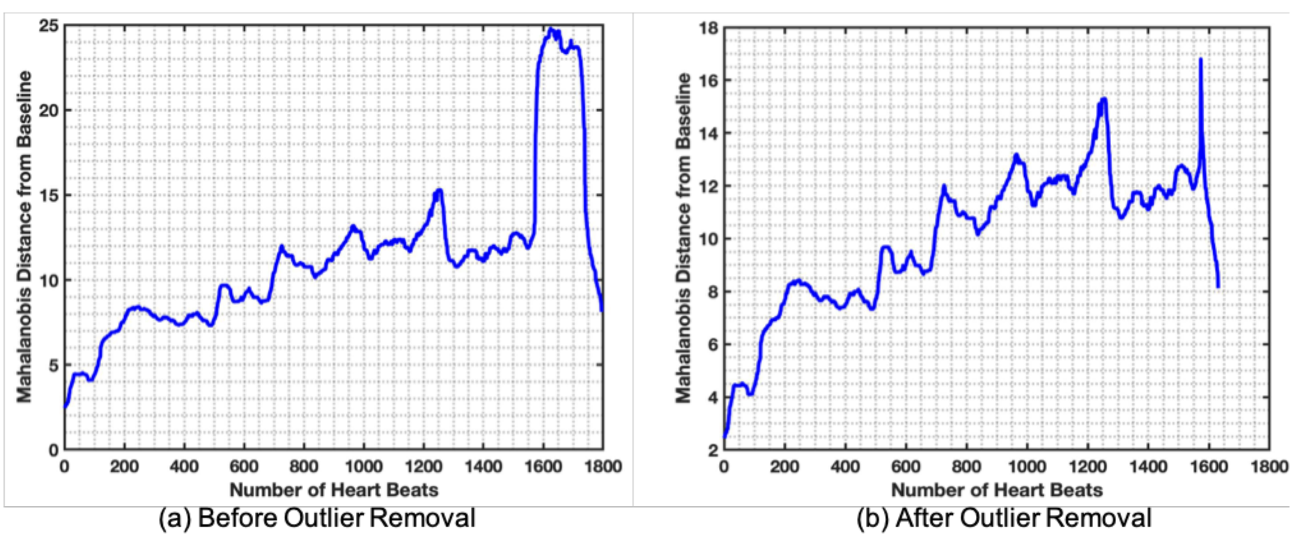
Supplemental Figure S1: Automatic artifact removal of ECG: (a) Filtered ECG and SCG signals. (b) ECG is passed through artifact detection algorithm. Segments with artifacts are chosen when the difference (E_d) of upper (E_u) and lower (E_l) envelope of the signal is outside the range of mean (μ) ± 3*standard deviation (σ) of E_d. The longest artifact free signal segment is chosen and the time stamp of the start and end of that segment is calculated. (c) Both ECG and SCG signals are segmented with the time stamps from the previous step, where the ECG was found to be artifact free.

Extracted SCG Features:

SCG features extracted were as follows: total band power (0-500 Hz), and band power in 3 Hz bandwidth frequency ranges from DC to 30 Hz. Additional frequency domain features were: the highest prominent frequency (fp1), second prominent frequency (fp2), and third prominent frequency in the power spectral density (PSD), and the amplitudes of the PSD at fp1, fp2, and fp3.

For prominent frequency, peaks in the PSD of the frame were ranked according to their amplitudes, and the highest, second highest and third highest amplitude were used to locate the first, second and third prominent frequency accordingly.

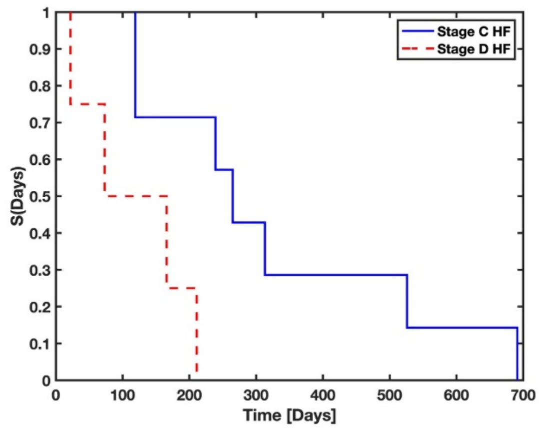
Outlier Removal using Mahalanobis Distance: Before training the regression model to estimate VO_2, we removed outliers in the SCG signals. For each subject, the first 50 averaged frames from the rest signal was assumed as baseline frames and all the features (for a particular feature set) of the 50 frames were averaged to create baseline feature distribution. The Mahalanobis distance (1) was calculated between the baseline feature distribution and each averaged heartbeat frame for a particular subject. The underlying hypothesis was that the wearable signal would change in morphology with various intensity of exercise and it would vary the most at peak exercise compared to baseline which would be reflected by the Mahalanobis distance. The first and third quartile (Q1 and Q3) were extracted and interquartile range (IQR) was calculated as $IQR = Q3 - Q1$, for subject-wise distribution. A particular frame was considered as outlier if the Mahalanobis distance of the frame is lower than $(Q1 - 1.5xIQR)$ or higher than $(Q3 + 1.5xIQR)$ of the distribution. These outlier frames and corresponding target variable were removed from the dataset. Figure S2 shows an example of the artifact removal for one subject. The distance calculated for each frame was used as a feature in the regression model, which makes the total number of features equal to *f+1*.



Supplemental Figure S2: Outlier removal using Mahalanobis distance: (a) Mahalanobis distance before outlier removal. (b) Mahalanobis distance after outlier removal.

Preprocessing Wearable Signal Features for Classifier: For each subject, ensemble averaged heartbeats were ranked from lowest to highest using corresponding Mahalanobis distance (described above) for a particular subject and highest 100 heartbeats were chosen for each subject for further classification analysis, giving us total of 4400 heartbeats from 44 subjects. Underlying hypothesis of choosing highest 100 heartbeats was that the subjects were classified based on the peak exercise capacity during CPX and wearable signal segments correspond to peak exercise capacity would reflect the difference between stage C and stage subjects. These heartbeats were labeled for each subject based on the true class for that particular subject.

Survival Analysis with Subsequent Events: We have been following the subjects for subsequent events: left ventricular assisted device (LVAD) implantation, heart transplant, and cardiovascular death, occurring six months following the initial collection of data. In the cases where one CPX subject had multiple events (e.g. LVAD, followed by transplant later), we used only the first occurring event for the survival analysis. Based on these events, Kaplan-Meier analysis based survival data was derived for the patients in stage C and D groups, as shown in Figure S3 below.

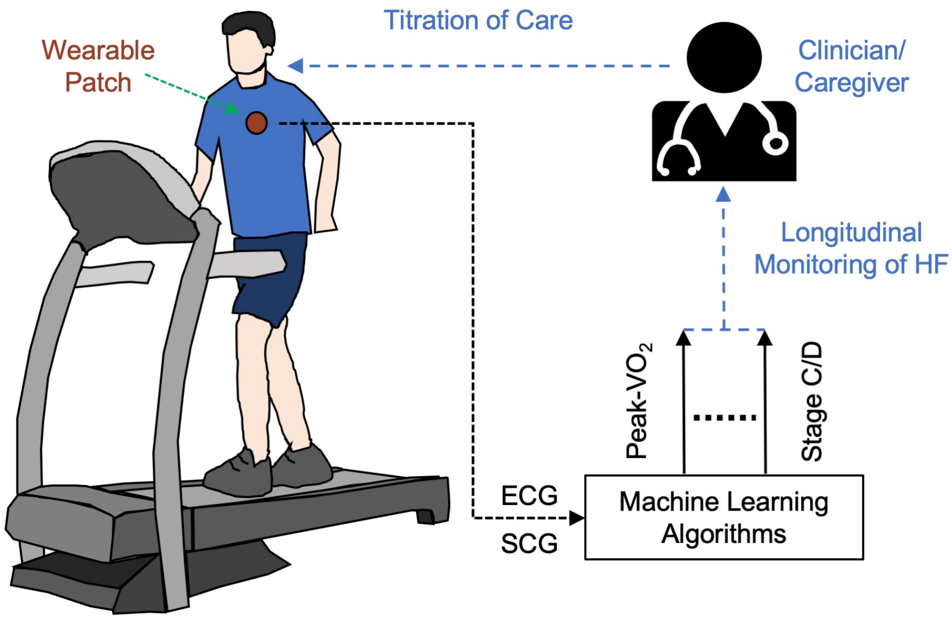


Supplemental Figure S3: Kaplan-Meier estimate of survival analysis between stage C and D HF.

Reference:

1. Mahalanobis PC. On the generalized distance in statistics. National Institute of Science of India, 1936.

Supplemental Figure 1: Concept of Longitudinal HF Management using a Wearable Patch



Supplemental Figure 1. Illustration of our envisioned three-step process for longitudinal monitoring of HF patients: 1) Recording of SCG and ECG signals using a custom-built wearable patch during exercise and daily activities. 2) Estimation of cardiopulmonary gas exchange variables and prediction of clinical state of HF (stage C or D). 3) Intervention by a clinician based on the longitudinal assessment of cardiopulmonary parameters and HF status (future work).