The Journal of Open Source Software

Wearablehrv: A Python package for the validation of heart rate and heart rate variability in wearables

Mohammadamin Sinichi[1,3¶], **Martin Gevonden**[2,3], and **Lydia Krabbendam**[1,3]

1 Department of Clinical, Neuro- & Developmental Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands **2** Department of Biological Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands **3** Institute Brain and Behaviour (iBBA), Amsterdam, The Netherlands ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Olivia Guest

Reviewers:

- @jamesheathers
- @AKuederle

Submitted: 10 November 2023
Published: unpublished

Summary

Wearable devices that monitor physiology have become ubiquitous, and include smart watches, smart jewelry, and smart textiles. The market for these devices is rapidly expanding with new brands and products. These devices measure a variety of signals, which are translated into a large amount of different features, of which heart rate (HR) and heart rate variability (HRV) are among the most common. These features are particularly interesting , not only for consumers but also to researchers, because they are predictive of mental and physical health outcomes, and easily obtained. However, for manufacturers there may be a trade-off between user acceptability on one hand and accuracy on the other, where profit margins typically turn out to be the decisive factor. Therefore the following question continually comes up: is the cardiac data recorded by this new watch/ring/shirt accurate enough to use in research?

The `wearablehrv` Python package offers a comprehensive pipeline for validating the accuracy of HR and HRV measurements. It allows for advanced statistical analyses on device agreement from beat-to-beat cardiac data. The package's graphical user interface facilitates pre-processing, visualization, and data analysis at both individual and group levels. As input, a user of `wearablehrv` should use a criterion device, preferably an electrocardiograph (ECG). Simultaneously acquired data from the device(s) of interest can then be compared against the criterion. The only required inputs for the pipeline are the inter-beat intervals (IBIs) and timestamps for each device; the rest is fully handled by the `wearablehrv` package.

Statement of Need

The use of wearables in psychophysiology and sports sciences has exponentially increased over the past decade. While the golden standard is electrocardiography, for prolonged monitoring many studies use devices which measure HR and HRV using Photoplethysmography (PPG). This technique optically measures the changes in blood volume in peripheral tissues (e.g., earlobe, wrist, arm, fingertip) and substitutes detected pulses for the actual contraction of cardiac ventricular muscles(Challoner & Ramsay, 1974). Although the PPG method is promising and versatile, several studies have highlighted concerns about its validity in identifying HR and HRV, especially when the user is in motion (Allen, 2007; Bent et al., 2020; Hill et al., 2015; Nederend et al., 2017; Pinheiro et al., 2016; Quintana et al., 2016; Schäfer & Vagedes, 2013; Stone et al., 2021).

The oversight of these wearables' validity may stem from the absence of an integrated and user-friendly way to assess the validity of novel wearables. A validation pipeline would guide users from raw data collection through pre-processing and advanced statistical analyses to arrive at the necessary parameters and visualizations necessary to assess device agreement. The `wearablehrv` package was developed to address this gap. Provided that a wearable device (either PPG or ECG) allows for export of the complete time series of recorded IBIs, this Python package makes it relatively easy to establish the validity of a novel wearable in just a few steps. In summary, `wearablehrv` is a Python package tailored for data preparation, pre-processing, feature extraction, comparison, visualization, and both individual and group statistical analyses of heart rate and heart rate variability metrics from wearable devices that transmit raw IBIs and timestamps. The inclusion of graphical user interfaces (GUI) in most functions grants researchers and wearable users the flexibility in validating an unlimited number of wearables across a range of conditions.

Main Features and Basic Usage

In this section, we provide an overview of the main functions and basic usage of the `wearablehrv` pipeline. For detailed documentation of all the functions and examples on data recording and initiating validation, please refer to README.md and documentation.ipynb.

The pipeline is divided into two parts: the individual pipeline, that focuses on the data from a single individual, and the group pipeline, which combines the data from all the observations in a single study.

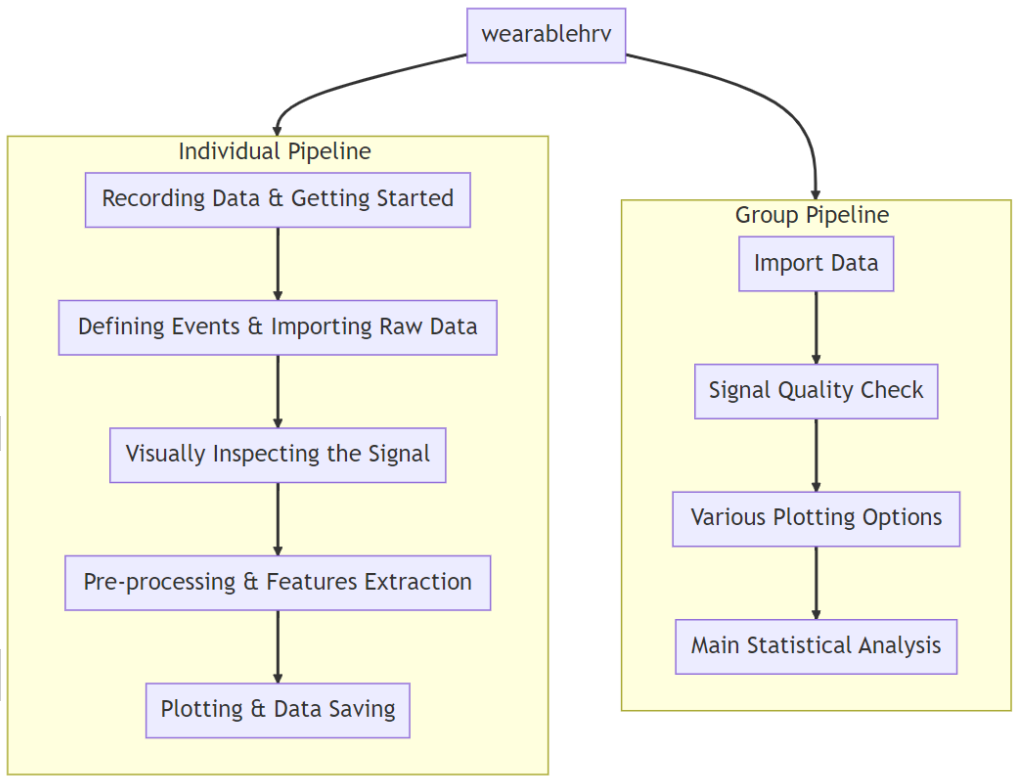

Figure 1: A summary of the the main functionalities of wearablehrv Python package.

The individual pipeline allows for the validation of a wearable of interest against a criterion device (e.g., an ECG) under unlimited conditions. Data necessary for the pipeline should be in a .csv file with UNIX timestamps (specified in milliseconds) and IBIs, in two separate columns. For defining events and importing recorded data, the pipeline can be initiated with the `define_events` function, either by referencing a pre-specified .csv file or by using the GUI. The `import_data` function is then used to bring the continuous time series from all devices into the pipeline. Data segmentation into smaller, condition-specific chunks is achieved with the chop_data function.

One of the primary strengths of wearablehrv is the visual_inspection function, which allows for simultaneous visualization of IBI signals and addresses the challenge of correcting devices lag when their internal clocks are not in sync. With the assistance of the GUI, adjusting for this lag becomes straightforward.

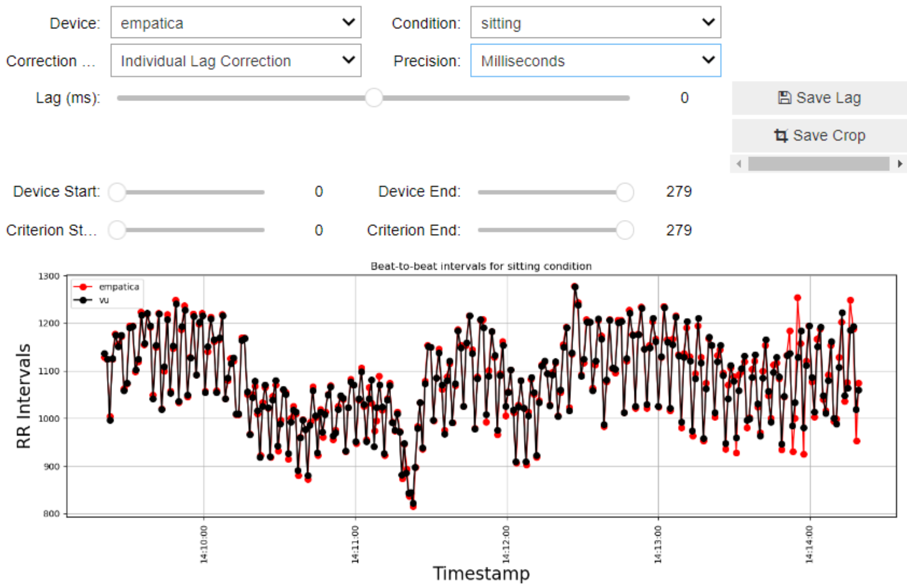

Figure 2: The black line shows the IBIs of the criterion device (ECG). The red line represents the IBIs of a given PPG device. By toggling the Device and Conditions widgets, one can easily explore other devices and conditions. The Lag slider allows for lag correction between the devices, and it is possible to crop a part of the signal if necessary.

Pre-processing and feature extraction are streamlined using the pre_processing function and the data_analysis function from the hrv-analysis python package(Robin Champseix, 2021), outputting time domain and frequency domain features for all devices and conditions.

To facilitate comparing the criterion device with the device of interest and visualization of results, several plotting functions are provided, including result_comparison_plot, bar_plot, line_plot, radar_plot, and unfolding_plot. All time and frequency features for every device and condition can be exported for later use in the Group Pipeline via the save_data function.

Moving to group-level analysis, the import_data function of the group module aggregates individual case analyses. A significant advancement in wearable validation is the ability to identify the amount of missing, poor and acceptable data in each device, and generate detailed reports. The signal_quality function allows for such signal quality assessment across devices and participants, providing detailed reports for informed decision-making.

For visualization across the aggregated dataset, the group module offers violin_plot, box_plot, radar_plot, hist_plot, matrix_plot, and density_plot.

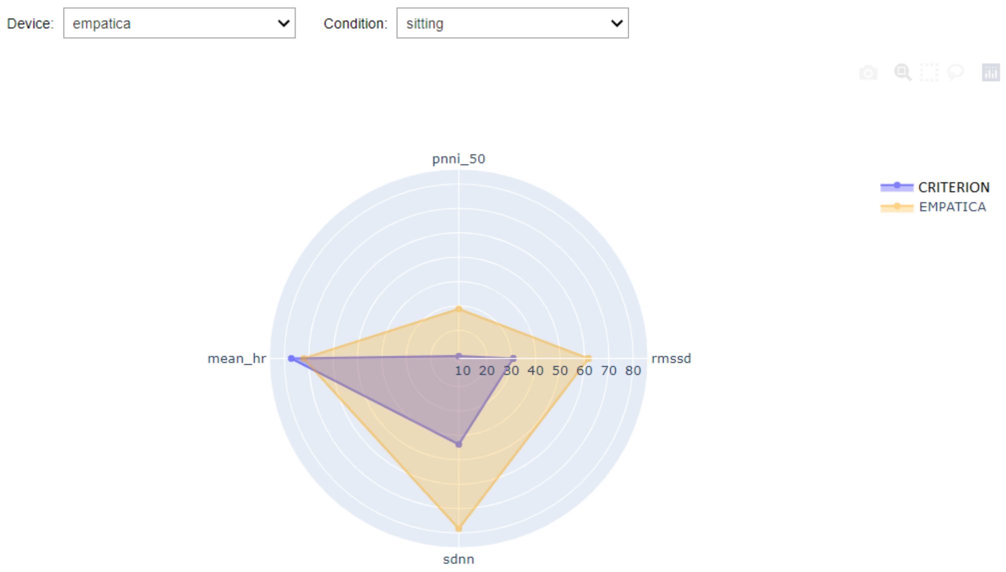

Figure 3: An example of the radar plot at the group level, illustrating a comparison between a criterion device and a specified device for `pnni_50`, `rmssd`, `sdnss` (all distinct extracted features for HRV), and mean heart rate (`mean_hr`). Utilizing the `Device` and `Condition` widgets facilitates easy transitioning between a multitude of devices and conditions.

The culmination of the group pipeline involves the main statistical analyses — regression analysis, ICC, and Bland-Altman analysis — enabled by the `regression_analysis`, `icc_analysis`, and `blandaltman_analysis` functions, respectively. These are complemented by their corresponding plotting functions, which can be used to make decisions about device validity, or included in a publication.

Acknowledgements

This study is an output of the ALIVE project (Improving Adolescent mentaL health by reducing the Impact of PoVErty). This work was supported by the Wellcome Trust [221940/Z/20/Z]. We acknowledge contribution from Dr. Denise J. van der Mee and advice from Dr. Marco Altini.

References

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, *28*(3), R1. https://doi.org/10.1088/0967-3334/28/3/R01

Bent, B., Goldstein, B. A., Kibbe, W. A., & Dunn, J. P. (2020). Investigating sources of inaccuracy in wearable optical heart rate sensors. *Npj Digital Medicine*, *3*(1), 1–9. https://doi.org/10.1038/s41746-020-0226-6

Challoner, A. V. J., & Ramsay, C. A. (1974). A photoelectric plethysmograph for the measurement of cutaneous blood flow. *Physics in Medicine & Biology*, *19*(3), 317. https://doi.org/10.1088/0031-9155/19/3/003

Hill, L. K., Hu, D. D., Koenig, J., Sollers, J. J., Kapuku, G., Wang, X., Snieder, H., & Thayer, J. F. (2015). Ethnic Differences in Resting Heart Rate Variability: A Systematic Review and Meta-Analysis. *Psychosomatic Medicine*, *77*(1), 16–25. https://doi.org/10.1097/PSY.0000000000000133

Nederend, I., Harkel, A. D. J. ten, Blom, N. A., Berntson, G. G., & Geus, E. J. C. de. (2017). Impedance cardiography in healthy children and children with congenital heart disease: Improving stroke volume assessment. *International Journal of Psychophysiology*, *120*, 136–147. https://doi.org/10.1016/j.ijpsycho.2017.07.015

Pinheiro, N., Couceiro, R., Henriques, J., Muehlsteff, J., Quintal, I., Goncalves, L., & Carvalho, P. (2016). Can PPG be used for HRV analysis? *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2016*, 2945–2949. https://doi.org/10.1109/EMBC.2016.7591347

Quintana, D. S., Alvares, G. A., & Heathers, J. a. J. (2016). Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): Recommendations to advance research communication. *Translational Psychiatry*, *6*(5), e803–e803. https://doi.org/10.1038/tp.2016.73

Robin Champseix, C. L. C., Laurent Ribiere. (2021). A python package for heart rate variability analysis and signal preprocessing. *Journal of Open Research Software*, *9*, 28. https://doi.org/10.5334/jors.305

Schäfer, A., & Vagedes, J. (2013). How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram. *International Journal of Cardiology*, *166*(1), 15–29. https://doi.org/10.1016/j.ijcard.2012.03.119

Stone, J. D., Ulman, H. K., Tran, K., Thompson, A. G., Halter, M. D., Ramadan, J. H., Stephenson, M., Finomore, V. S., Galster, S. M., Rezai, A. R., & Hagen, J. A. (2021). Assessing the Accuracy of Popular Commercial Technologies That Measure Resting Heart Rate and Heart Rate Variability. *Frontiers in Sports and Active Living*, *3*. https://doi.org/10.3389/fspor.2021.585870