# Tell me how I feel! Development of an algorithm to predict depression symptoms based on wearable data

Hammelrath L. [1], Heinrich M.[1], Burggraf P.[2], Schenk, H. [2] , Knaevelsrud C. [1]

[1] Division of Clinical Psychological Intervention, Faculty of Psychology, Freie Universität Berlin, Germany, [2] Thryve GmbH, Berlin, Germany

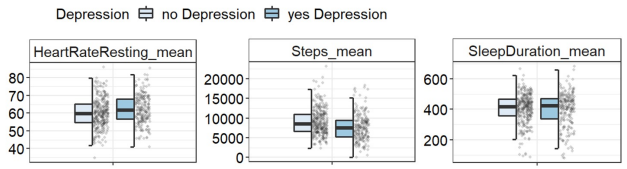
## Background

- Timely detection of depression and treatment initiation is important to reduce the risk of future episodes
- Wearables might significantly contribute here, as they can provide objective and low-treshold data from the individuals‘ everyday life
- Previous studies found promising correlations of sensor-based information on sleep quality and physical activity with depression[1]
- Wearable AI - the combination of remote sensing data with machine learning (ML) methods – is a promising approach to develop predictive models based on a large and potentially noisy databases
- A recent meta-analysis[2] revealed that Wearable AI can detect a current depression with a good, but not optimal accuracy of 70-89%. They concluded, that self-report items should be used in conjunction with wearable data
- Most included studies, however, used small sample sizes
- We aim to predict current depression status based on (a) wearable data only, (b) self-report data only or (c) a combination of both using state of the art ML methods on a comprehensive sample

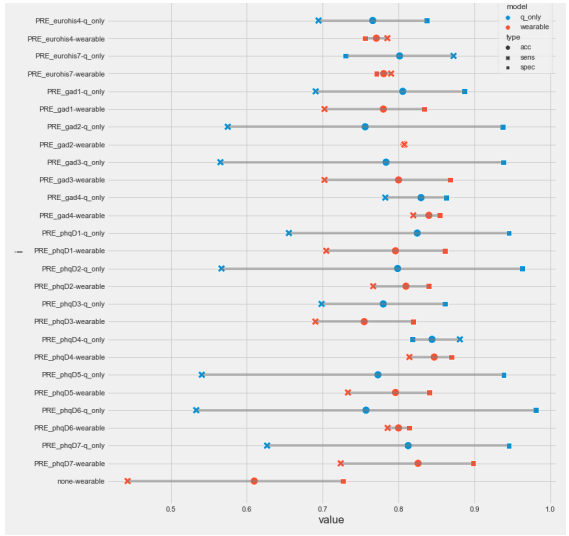
## Database

- The **Blueswatch** study is a cooperation of *Freie Universität Berlin*, *Thryve* and the *Techniker Krankenkasse*
- Participants took part with their own wearable and provide remote sensing data for 4 weeks after they filled out self-reports on mood (PHQ-8), anxiety (GAD), and quality of ife (EUROHIS)
- Our database encompasses n= 412 participants providing 14 days of remote sensing data; For our analyses we used:
  - Sleep duration (daily mean + SD)
  - Step count (daily mean + SD)
  - Heart rate at rest (daily mean + SD)
- Around half of patients filled out self-reports before sharing their wearable data

## Results




**Outcomes per model type**




**Blueswatch Is ongoing** Take part!

Literature and Supplementary Information

## Analysis

- Outcome: binary, defining the presence of depression as a PHQ-8 sum score >= 10
- 5-fold cross-validation with a train-test split of 80/20; downsampling of the majority class
- Training Features encompassed:
  - Sleep duration (M + SD), averaged across 14 days
  - Step count (M + SD), averaged across 14 days
  - Heart rate at rest (M + SD), averages across 14 days
  - Self report: age, gender, PHQ single items, GAD single items, EUROHIS single items
- Preprocessing: Features were scaled (MinMax Scaling) or one-hot-encoded
- ML algorithm: **LightGBM** classifier
  - Gradient boosting model; uses tree bases learning algorithms
  - Boosting type = „dart“
  - Standart hyperparameters

## Discussion

- To achieve sufficient predictive power the integration of self-report items was necessary; wearables only lead to an accuracy of 63%
- Our results match with the meta-analysis by Abd-Alrazaq et al. (2023)
- Limitations:
  - Preprocessing was basic. In following studies we might use standardized and/or normed values
  - Try out other models that are more adequate for the database
  - Some participants provided wearable data after their self-report; next analyses should include pre-outcome data only
- Outlook: data collection is ongoing. As a next step, we plan to implement an updated version of our algorithm using self-report + wearable data into an app for people insured at *Techniker Krankenkasse*

Supported by Deutsche Forschungsgemeinschaft – FOR 5187 (project number 442075332)