Stress in Action Wearables Database: A database of non-invasive wearable monitors with systematic technical, reliability, validity, and usability information

Melisa Saygin[1]†; Myrte Schoenmakers[1]†; Magdalena Sikora[2]†; Thomas Vaessen[2], Matthijs Noordzij[2], Eco de Geus[1]

[1]: Vrije Universiteit Amsterdam
[2]: University of Twente
†: These authors contributed equally to this work.

Why needed?

- Rapid growth of available consumer and research-oriented wearables
- To systematically choose a wearable among the multitude of available devices, one should perform:
  - Get a detailed overview of technical and other device specific information
  - Retrieve synopsis of Reliability, Validity and Usability studies per wearable

To reduce the time and effort required to choose the appropriate wearable, the Stress in Action Wearables Database provides this detailed overview.

To the DATABASE

To the PAPER

Results

- 54 wearables were included in the first version of the database, and 172 are listed to be added in the future iterations.
- Among the wearables included in the first version, 64.8% record PPG, 50.0% ECG, 35.2% record Skin temperature, 31.5% record blood oxygen saturation level (SpO2) and 31.5% can record EDA. The respiration signal, either derived from ICG or respiration belts, is measured by only 5.6% of the devices.
  - Wearables were intended for Consumers (35), followed by Research (18) and Clinical use (10), including nine devices intended for multiple user fields.

The Stress in Action Wearables Database facilitates rapid and systematic wearable selection for stress research

A section of the Stress in Action Wearables Database

showing 11 of the 53 columns for six devices

DEVICE	ESTIMATED COSTS (€)	WEARABLE TYPE	PPG	ECG	RAW DATA AVAILABLE	DEVICE STORAGE CAPACITY (HOURS)	BATTERY LIFE (HOURS)	HIGHEST LEVEL OF VALIDATION; NUMBER OF PAPERS	GENERAL VALIDITY CONCLUSION
EmbracePlus 	1.984; details	Watch	1; reflection; 26-208; wrist; green & red & infrared	0	1	24	48	Internal; 1	Reliable and valid for use in clinical SpO2 measurement under static conditions, for both male and female subjects as well as for different skin tones.
VU-AMS Core 	5995; details	Box + electrodes	0	1; 3; 1000; electrode; lead I; lead II; aVF	1	30	28	No validation; 0	No validation studies yet- follow up of VU-AMS 5fs device.
Fitbit Charge 5 	150	Watch	1; reflection; NP; wrist; red and infrared	0	0	168	168	External; 1	Good HR detection for sinus rhythm, poor for atrium fibrillation. Irregular rhythm notification acceptable for guiding decision regarding atrial fibrillation patients
Oura Ring Gen3 	314; details	Ring	1; reflection, transmission; 250; finger; green & red & infrared	0	0	168	168	External; 3	High validity for HR and RMSSD and acceptable accuracy for nocturnal AVNN, pNN50, HF and SDNN in average-per-night tests with high error rates for LF and LF:HF ratio. Participant compliance was high over several weeks.
Hexoskin Pro 	1.984; details	Box + vest	0	1; 1; 256; touch; lead I	1	2400	36	External; 24	RR and HR were found to be valid with respect to good quality reference devices and also in ecological settings, but a mix of positive and negative results exist about the validity of energy expenditure and minute ventilation.
Research Ring 	2020; details	Ring	1; reflection; 400; finger; green	1; 1; 800; touch; lead I, lead II	1	Bluetooth	10	No validation; 0	No validation studies yet- device is quite new.

Methods – data retrieval

DEVICE SPECIFIC INFORMATION

Information – obtained from the device manual – about the device is categorized as follows:

Category	Examples
General device information	Wearable type, size, weight, costs
Signals	PPG, ECG, EDA, BP, accelerometer
Technical information	Battery life, water resistance
Data access	Raw data, parameters, software, storage
Reliability, validity and usability	Number of studies, general synthesis

RELIABILITY, VALIDITY AND USABILITY

Original experimental studies search using the following string:
"((Device Name) AND (valid* OR reliab* OR compar* OR accur* OR verif* OR usab* OR "user experience" OR "user friend*" OR user-friend*)"

The following databases were searched: **PubMed, Web of Science, Scopus, IEEExplore, ACM Digital Library**

Inclusion Criteria	Exclusion Criteria
• Assessment of parameter level convergent validity, test-retest reliability, and/or usability.	• Studies on construct validity only
• Assessment of convergent validity	• Studies on machine learning based detection of secondary outcomes
• Peer-reviewed articles and conference proceedings published in English	• Meta-analyses and reviews
	• Theses, grey literature, other text that was not peer reviewed

Methods – literature search

ASREVIEW PROCESS

Phase 0: Compilation of search results

If there are >100 papers, proceed to phase 1, otherwise review all papers

Phase 1: Train the ASReview algorithm on a subset of the papers

(1) Screen a random subset of at least 10% of the total number of papers

(2) Calculate the ratio of relevant to non-relevant papers

(3) Apply that ratio to the whole dataset to arrive at the expected number of relevant records in the total dataset (*RR_T*)

Phase 2: Apply the trained ASReview algorithm on the remaining papers

Accept the papers for further review until:
(1) the visual inspection of the recall plot clearly shows a plateau; *AND*
(2) at least twice the *RR_T* records have been screened; *AND*
(3) a minimum of 35% of the total dataset has been screened; *AND*
(4) no extra relevant papers have been identified in the last 50 papers

OR

Review all the papers

Scoring the wearables for short- and long-term use

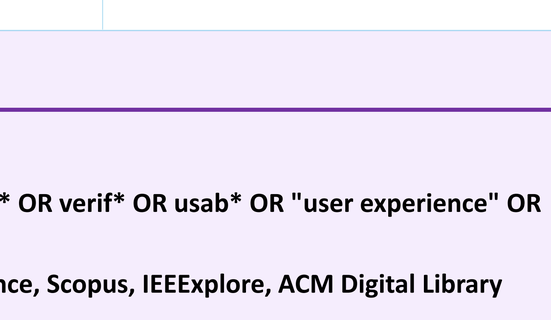
Long-term importance of criteria	
• GDPR approval	High
• CE approval	High
• Price	High
• Number of different physiological parameters that can be measured by the device (e.g., HR, Skin temperature)	High
• Wearable type, location, and weight and size	High
• Usability outcome (if available)	High
• Battery life	High
• All validity-reliability criteria	High/medium
• Parameter sampling window	High/medium
• Data storage capacity	High/medium
• Data transfer method	Medium
• Data Storage Method	Medium
• Charging duration	Medium
• Charging method	Medium
• Raw data availability	Medium
• Data Transfer Compatibility	Medium
• Waterproof	Medium/low
• Bio-cueing	Medium/low
• Biofeedback	Medium/ow
• FDA approval	Low

Short-term importance of criteria	
• GDPR approval	High
• CE approval	High
• All reliability– validity criteria	High
• Number and type of different signals that can be measured by the device (e.g., PPG, ECG, accelerometer)	High
• Raw data availability	High
• Provided parameters	Medium
• Parameter sampling window	Medium
• Price	Medium
• Wearable type, location, and weight and size	Medium
• Usability outcome (if available)	Medium
• Data storage capacity	Medium/low
• Data transfer method	Medium/low
• Data storage method	Medium/low
• Battery life	Low
• Charging duration	Low
• Charging method	Low
• FDA approval	Low
• Waterproof	Low
• Bio-cueing	Low
• Biofeedback	Low

- Different criteria were rank-ordered in importance by the curators for scoring a device for short-term (2-days) and long-term (2+ weeks) research use
- Following this list of criteria, each device was independently scored from 0 to 10 by the three first co-authors who were blinded to each other's scores. The scores were averaged, making the short- and long-term "SiA expert scores"
- High interrater reliability across the raters was obtained for both short-term, .87, 95% CI = [.78, .92], $F(50, 100) = 8.0$, $p < .001$. ,and long-term use, .85, 95% CI = [.76, .91], $F(50, 100) = 6.6$, $p < .001$.

The Stress in Action Wearables Database is a living document

- In contrast to the wearable inventories or review studies with a static document, the SiA wearables database will be continuously updated every 6 months. This enables the entry of newly released wearables and updating of the validity, reliability, and usability information on existing wearables.



The wearables database will soon also be available with a user-friendly interface with filtering options!
Stay tuned at **stressinaction.nl**

STRESS IN ACTION