Investigation of the use of a sensor bracelet for the presymptomatic detection of changes in physiological parameters related to COVID-19: an interim analysis of a prospective cohort study (COVI-GAPP)

Martin Risch,1,2,3 Kirsten Grossmann,4,1,4 Stefanie Aeschbacher,5 Ornella C Weideli,6,7 Marc Kovac,3 Fiona Pereira,8,9 Nadia Wohlwend,3 Corina Risch,3 Dorothea Hillmann,3 Thomas Lung,3 Harald Renz,10,7 Raphael Twerenbold5,8 Martina Rothenbühler9,9 Daniel Leibovitz,9 Vladimir Kovacevic,9 Andjela Markovic9,10,11 Paul Klaver,12 Timo B Brakenhoff,12 Billy Franks,12 Marianna Mitrata,13,14 George S Downward10,13,14 Ariel Dowling15, Santiago Montes,16 Diederick E Grobbee,13,14 Maureen Cronin,9 David Conen,17 Brianna M Goodale,9 Lorenz Risch,1,3,18 on behalf of the COVID-19 remote early detection (COVID-RED) consortium


ABSTRACT

Objectives We investigated machine-learning-based identification of presymptomatic COVID-19 and detection of infection-related changes in physiology using a wearable device.

Design Interim analysis of a prospective cohort study.

Setting, participants and interventions Participants from a national cohort study in Liechtenstein were included. Nightly they wore the Ava-bracelet that measured respiratory rate (RR), heart rate (HR), HR variability (HRV), wrist-skin temperature (WST) and skin perfusion. SARS-CoV-2 infection was diagnosed by molecular and/or serological assays.

Results A total of 1.5 million hours of physiological data were recorded from 1163 participants (mean age 44±5.5 years). COVID-19 was confirmed in 127 participants of which, 66 (52%) had worn their device from baseline to symptom onset (SO) and were included in this analysis. Multi-level modelling revealed significant changes in five (RR, HR, HRV, HRV ratio and WST) device-measured physiological parameters during the incubation, presymptomatic, symptomatic and recovery periods of COVID-19 compared with baseline. The training set represented an 8-day long instance extracted from day 10 to day 2 before SO. The training set consisted of 40 days measurements from 66 participants. Based on a random split, the test set included 30% of participants and 70% were selected for the training set. The developed long short-term memory (LSTM) based recurrent neural network (RNN) algorithm had a recall (sensitivity) of 0.73 in the training set and 0.68 in the testing set when detecting COVID-19 up to 2 days prior to SO.

Strengths and limitations of this study

⇒ Large sample size from a well-characterised and healthy national cohort.
⇒ Wearable device technology combined with machine learning to monitor health parameters related to early detection of COVID-19 infections.
⇒ Solely data from laboratory confirmed COVID-19 infections were used.
⇒ Data from one single study centre may limit the generalisability of our findings.
⇒ Small subsample of COVID-19 positive cases with sufficient high-quality data.

Conclusion Wearable sensor technology can enable COVID-19 detection during the presymptomatic period. Our proposed RNN algorithm identified 68% of COVID-19 positive participants 2 days prior to SO and will be further trained and validated in a randomised, single-blinded, two-period, two-sequence crossover trial.

Trial registration number ISRCTN91255782; Pre-results.

INTRODUCTION

One of the primary ways of controlling the spread of SARS-CoV-2 involves identification, tracing and isolation programmes implemented in several countries.1 With multiple SARS-CoV-2 variant strains emerging, countries have prioritised vaccine rollouts, searches for alternatives to quarantine and
identification of individuals with COVID-19. Reverse transcription-polymerase chain reaction (RT-PCR), serological testing, surveys, temperature measurements and symptom checks are used to detect COVID-19. However, these methods are usually unable to identify pre-symptomatic or asymptomatic individuals.

Recent studies have highlighted the need to identify potential cases prior to symptom onset (SO) to prevent virus transmission. Asymptomatic patients are likely to ignore safety precautions, leading to increased virus transmission. Detection of COVID-19 during the asymptomatic or pre-symptomatic stage facilitates early isolation, thereby limiting contact with susceptible individuals. Commonly reported COVID-19 symptoms include fever, coughing, chest tightness, difficulty breathing, fatigue, dyspnoea, myalgia, sputum production, headache and gastrointestinal symptoms. While molecular tests are continuously used to confirm infections, the logistics and costs of repeat tests across populations are prohibitive. Recently, scientists have called for further research investigating whether wearable medical devices such as Ava-bracelets and direct-to-consumer products such as Fitbit, smartwatches and other activity trackers could be used for such surveillance.

Here, we assess the use of an existing regulated wearable medical device (Ava-bracelet) to analyse COVID-19-related changes in various physiological parameters across four infection-related periods: incubation, pre-symptomatic, symptomatic and recovery. To our knowledge, this is the first prospective study to measure physiological changes in respiratory rate (RR), heart rate (HR), HR variability (HRV), wrist-skin temperature (WST) and skin perfusion to develop an algorithm to detect pre-symptomatic COVID-19 infection.

METHODS
Study design and participants
Participants from the ongoing observational population-based prospective cohort study (Genetic and Phenotypic Determinants of Blood Pressure and Other Cardiovascular Risk Factors (GAPP); n=2170) in the Principality of Liechtenstein were invited to participate in the current study (COVI-GAPP). Active since 2010, the GAPP study was designed to understand the development of cardiovascular risk factors in the general population better (ie, healthy adults aged 25–41 years). The exclusion criterion regarding participation in the COVI-GAPP study was individuals who did not provide written informed consent. The first COVI-GAPP participants were enrolled in April 2020, and the data used for this interim analysis was collected through March 2021 (n=1163). This COVI-GAPP interim analysis was preplanned as a pilot study to provide an initial algorithm for the COVID-RED project (n=20 000), a randomised, single-blinded, two-period, two-sequence crossover trial.

Bracelet, app and participant compliance
The Ava-bracelet (version 2.0; Ava AG, Zurich, Switzerland) is an FDA-cleared and CE-certified fertility aid bracelet that complies with international regulatory requirements and applicable standards. The wrist-worn tracker is commercially available at US$ 279 and consists of three sensors that measure five physiological parameters simultaneously: RR (breaths per minute), HR (beats per minute), HRV (ms), WST (°C) and skin perfusion (online supplemental figure S1). Although the Ava-bracelet measures multiple forms of HRV, we focused on two time-dependent and one frequency-dependent measurements: SD of the normal-to-normal interval (SDNN), root mean square of successive differences (RMSSD) and HRV ratio (see online supplemental material). In addition to the physiological parameters of interest, the Ava-bracelet measures sleep quantity (duration) and sleep quality using a built-in accelerometer. Prior studies have demonstrated how device data can inform a machine-learning algorithm that detects ovulating women’s most fertile days in real time with 90% accuracy. Worn only while asleep, the Ava-bracelet saves data every 10 s and requires at least 4 hours of relatively uninterrupted sleep. The participants synchronised their bracelets with a complementary smartphone app on waking, transferring data from the device to Ava’s backend system.

Although no study-specific adjustments were applied to the hardware of the Ava-bracelet, the complementary app had a customised user functionality developed by the manufacturer specifically for the COVI-GAPP study. Participants could still see and monitor changes in the physiological parameters in the app; however, they did not receive messages or algorithm-driven interpretations of their data (figure 1A). Participants recorded behaviours that may have interfered with the physiological parameters of interest (eg, alcohol, medication and drug intake), as such substances can alter central nervous system functioning (figure 1B). The daily diary in the custom app enabled participants to record COVID-19-related symptoms (figure 1C). To ensure the highest quality data, the study team reviewed a weekly compliance log that indicated which participants had synced their Ava-bracelets with the app during the preceding week. The study team followed up with the participants individually to mitigate operational challenges or log in issues.

SARS-CoV-2 antibody testing and RT-PCR testing
SARS-CoV-2 antibody tests were assessed at baseline (starting April 2020) and during follow-up (starting December 2020) by the medical laboratory Dr Risch Ostschweiz AG (Buchs SG, Switzerland). The tests were assessed with an orthogonal test algorithm that employed electrochemiluminescence assays. These assay test for pan-immunoglobulins directed against the N antigen and the receptor-binding domain of the SARS-CoV-2 spike protein. Seroconversion was assumed if the first blood sample was negative for SARS-CoV-2 antibodies and the second sample was positive.

If participants had any symptoms during the study period, they were encouraged to visit the Liechtenstein
National Testing Facility for RT-PCR testing. The testing facility was open daily allowing for higher testing frequencies than that in other European countries. RT-PCR was performed on either the COBAS 6800 platform (Roche Diagnostics, Rotkreuz, Switzerland) or the TaqPath assay on a QuantStudio 5 platform (Thermo Fisher Scientific, Allschwil, Switzerland). Participants diagnosed with COVID-19 contacted the study team to discuss their symptoms and health statuses. Additionally, participants provided their date of SO and overall symptom duration, enabling us to calculate the symptom end (SE) date.

**Questionnaires**

For the second antibody test, all participants were asked to complete a questionnaire providing personal information (age, sex), smoking status (current, past, never), blood group (A, B, AB, 0, unknown), number of children, exposure to household contacts who tested positive for COVID-19, working with people who have tested positive for COVID-19, and vaccination status. We calculated the body mass index (BMI) based on the height and weight collected from the GAPP database.

**Statistical analysis**

The primary objective was to determine whether different physiological parameters deviated from the baseline during COVID-19 infection. This information was used to develop a model for predicting COVID-19 infection before SO. To evaluate whether RR, HR, HRV, WST and skin perfusion deviated from baseline measurements during the four infection-related periods, we categorised the daily parameter measurements as occurring at baseline if the day (d) was >10 days prior to SO (ie, d>SO-10), the incubation period as SO-10 ≤ d < SO-2, and the presymptomatic period as SO-2 ≤ d < SO. We chose a cut-off of −2 days based on previous reports of infected participants becoming contagious 2 days before SO. Because the participants’ reported symptom durations varied, the measurements were categorised into the symptomatic infection category if SO ≤ d ≤ SE. Finally, the parameters collected after SE were classified as being in the recovery period (d>SE).

**Development of a machine-learning algorithm for detecting presymptomatic COVID-19 infection**

We chose a recurrent neural network (RNN) with long short-term memory (LSTM) cells for the binary classification of an individual as healthy or infected (positive for COVID-19) on a given day. LSTM networks have proven to be highly accurate in recognising time series patterns and events across large datasets. The internal structure
of such networks can memorise states and easily fetch or activate them, even if they were created many epochs ago. The LSTM network we implemented consisted of two hidden layers with 16 and 64 cells (figure 2). Its output activation was a sigmoid function, whereas the recurrent activation was a hyperbolic tangent (tanh) function. The output was limited to a range between 0 and 1 to ensure that the model yielded an overall probability of infection on a given day. A potential COVID-19 infection was indicated when this probability exceeded 0.5.

1. Data processing and multilevel model specification

All data processing and analyses were performed in R (version 3.6.1) and Python (version 3.6). Preprocessing of the data was performed to remove potential artefacts and ensure consistency with best practices25 (see online supplemental materials for detailed description). Further, we ran a series of multilevel models with random intercepts and slopes to determine the differences in physiological parameters during the infection-related periods compared with baseline. Given our continuous criterion, we modelled our outcomes of interest using residual maximum likelihood estimation and Satterthwaite df. Four binary variables were created, indicating the infection period to which a given measurement belonged (1=belonging to that period, 0=not belonging to that period). The reference baseline-period measurements were encoded as zero across all four binary variables. The reported results included unstandardised regression coefficients for each effect. When multiple models were possible for the same parameter, we chose the model using the percentile of the data (stable maxima) with the best fit (see online supplemental materials). To ensure a family-wise alpha level less than or equal to 0.05, we implemented Bonferroni correction for the seven analysed parameters (corrected alpha level of $p=0.007$) and adjusted the definition of marginal significance accordingly (i.e., $0.007 \leq p \leq 0.05$).

2. Data preparation and feature extraction for algorithm development

The Ava-bracelet records over a million data points per use. Therefore, we first identified the features that are most predictive of COVID-19. We normalised the night-time WST, RR and HR values to prime our model to detect deviations from baseline measurements and ensure greater stability in the measurements (e.g., to minimise interparticipant variability). Next, we compared the predictive performance of the raw features before engineering the novel composite features. We conducted a principal component analysis decomposition to test the correlation between the day of SO and other binary-labelled features (e.g., alcohol consumption). We also examined the correlation between WST and other physiological parameters to determine the potential autocorrelation prior to the model specification.

3. Training process

![Figure 2](image-url)
To limit our analysis to symptomatic COVID-19 cases, participants had to report the date of SO and record at least 28 days of bracelet data prior to that date. The full 4 weeks of data were required to ensure accurate baseline readings and enable the algorithm to account for cyclical variations in parameters attributable to monthly hormonal changes. Thus, each participant included in the analysis had at least 29 consecutive days of data recorded using the bracelet. We partitioned the data into 8-day sequences, enabling the algorithm to compare the physiological parameters across 8-day windows. This means that each user had more negative (class 0; ‘healthy’ days) sequences in the distribution than positive sequences (class 1; ‘infected’ days (eg, SO-10 to SO-2) as shown in figure 3).

We selected a binary cross-entropy loss function for the RNN by using a stochastic gradient descent (SGD) optimiser. Owing to the sample size, we set the learning rate to 0.007 and 2000 epochs, while also enabling an early stopping mechanism to prevent model overfitting. We trained our RNN 10 times, randomly splitting our sample into a training set (70% of participants) and a test set (30% of participants) for each instance. We report the metrics of the best-performing RNN model selected according to the following recall equation:

\[
\text{overall_recall} = (\text{recall_class_1_train} + \text{recall_class_0_train}) \times 0.7 + (\text{recall_class_1_test} + \text{recall_class_0_test}) \times 0.3 / 2
\]

Finally, because of the number of COVID-19 cases compared with healthy days in our dataset, we upsampled instances of class one through duplication, such that it was represented in our training set 1.15 times more than a given negative sequence (ie, class 0). Thus, the SGD optimiser treated the two classes as roughly equal and no longer overweighted the importance of the parameters predicting a healthy 8-day period. By training this LSTM model, we sought to leverage deep learning to predict the presymptomatic onset of COVID-19.

**RESULTS**

**Participants**

A total of 1163 participants (mean age=44.1 years, SD=5.6; 667 (57%) females) were enrolled in the COVI-GAPP study (figure 4). Of these participants, 127 (10.9%; 95% CI (9.3 to 12.8)) contracted COVID-19 during the study period. Ten infected participants were hospitalised for short-term monitoring, with breathing difficulties and fever as the main reported symptoms. Three asymptomatic infected participants were retrospectively identified using antibody tests. As seen in table 1, there were no differences in the sex ratio, age, BMI or smoking status between individuals who did or did not test positive for COVID-19 during follow-up (all \( \geq 0.30 \)). A significantly higher proportion of participants who contracted COVID-19 reported household contacts (n=58 of 1036 seronegative participants vs 53 of 127 seropositive participants; \( p<0.001 \)) or work colleagues who also had COVID-19 (n=230 of 1036 seronegative participants vs 49 of 127 seropositive participants; \( p<0.001 \)). On average, COVI-GAPP participants wore the Ava-bracelet for 1370.8 hours over the course of the study (SD=802.7), for a total of 1 453 006 hours. Of the 127 participants who tested positive for COVID-19, either through RT-PCR and SARS-CoV-2 antibody tests or antibody tests only, 66 users had worn their bracelet at least 29 days prior to SO which enabled sufficient data quality. Among these 66 participants, COVID-19 infection was confirmed by RT-PCR test and SARS-CoV-2 antibody test (n=48) or solely by antibody test (n=18).

1. Participants with confirmed COVID-19

Table 2 shows the clinical characteristics of COVID-19 positive participants, stratified according to their compliance with wearing the Ava-bracelet prior to SO. A series of 26 analyses of variance and chi-square tests with Bonferroni correction revealed that only BMI varied significantly between the two groups; noncompliant participants had a higher mean BMI (25.8 kg/m\(^2\), SD=4.0) than their compliant peers (23.8 kg/m\(^2\), SD=3.7; \( F(1, 116)=10.39, p=0.002 \)).

2. Compliant participants with confirmed COVID-19
Among the 66 compliant participants with COVID-19, 13,248 nights of data were collected (mean duration=200 nights, SD=47; range 72–284 nights) for a total of 124,079 hours (mean hours per participant=1880, SD=461.8). The compliant participants had a mean age of 42.9 years (SD=5.6) and most had never smoked (n=57; 86%). Their COVID-19 symptoms lasted for an average of 8.5 days (SD=5.0; range 1–25 days). Table 2 shows the frequency of the self-reported symptoms.

Physiological changes during the clinical course of COVID-19

Employing multilevel modelling, we observed significant changes in five (RR, HR, HRV, HRV ratio and WST) of the seven device-measured physiological parameters during the incubation, presymptomatic, symptomatic and recovery periods of COVID-19, compared with baseline. Table 3 lists the unstandardised coefficient values for each statistical model. The complete course of the different physiological parameters is shown in figure 5.

Respiration rate

COVID-19 positive participants had a significantly higher RR during the symptomatic period than at baseline (β_intercept = 15.1 breaths/min, SE=0.26; p<0.0001). Controlling for intraindividual variance, the nightly RR increased by 1.0 breaths/min (SE=0.18; p<0.0001). There were no significant differences in the RR detected between the baseline and other periods (all p≥0.114).

Heart rate

At baseline, the participants had a resting nightly HR of 55.4 beats per minute (bpm; SE=0.83; p<0.0001). During the incubation period, individuals’ HR increased significantly by 0.87 bpm (SE=0.29; p=0.004). HR remained elevated in the presymptomatic period, expected to be 1.0 bpm higher than that at baseline (SE=0.36, p=0.007). HR continued to increase following SO, beating 2.2 bpm faster than at baseline (SE=0.48, p<0.0001). Finally, even

Figure 4 Study flow chart. From 2170 GAPP participants, 1163 participants were enrolled in the COVI-GAPP study. A total of 127 participants presented laboratory-confirmed COVID-19 disease and from these, a total of 66 positive tested participants had complete bracelet data available used for the algorithm development.
after SE, participants had a significantly elevated HR (+0.87 bpm higher than baseline; SE=0.22, p=0.0002).

HRV: SD of the NN interval
Compared with a baseline SDNN of 59.6 ms (SE=1.4, p<0.0001), participants had significantly decreased SDNN in the incubation (βincubation = −1.5 ms, SE=0.59, p=0.0149), presymptomatic (βpre-symptomatic = −1.7 ms, SE=64; p=0.0086) and symptomatic (βsymptomatic = −1.4 ms, SE=0.73; p=0.0499) periods. Following SE, SDNN returned to baseline levels (βrecovery = −0.9 ms, SE=0.51, p=0.0787).

HRV: root mean square of successive differences
Our analyses did not reveal any significant phase-based differences in RMSSD for COVID-19 positive participants during their infection (all p≥0.5474) compared with baseline (βintercept = 43.7 ms, SE=1.2; p<0.0001).

HRV ratio
As with SDNN, multilevel analysis revealed a marginally significant decrease in HRV ratio during the incubation (βincubation = −0.01, SE=0.0; p<0.0001) and presymptomatic periods (βpre-symptomatic = −0.02, SE = −0.01; p=0.0165) compared with baseline (βintercept = 0.50, SE=0.02; p<0.0001). No significant difference in HRV ratio emerged between baseline and the symptomatic or recovery periods (all p≥0.5474).

Wrist skin temperature
Over and above participant level variance, WST increased by 0.13°C (SE=0.04; p=0.0001), 0.18°C (SE=0.05; p=0.001) and 0.3°C (SE=0.05; p<0.0001) during the incubation, presymptomatic and symptomatic periods, respectively, compared with baseline (βintercept = 35.3°C, SE=0.06; p<0.0001). WST remained elevated by 0.2°C relative to baseline, even during the recovery period (SE=0.03; p<0.0001).

Skin perfusion
No changes in skin perfusion were observed when comparing measurements during infection (all p≥0.5474) with baseline values (βintercept = −0.01, SE=0.0; p<0.0001).

Model specification and algorithm performance
The best-performing RNN consisted of composite features derived from the maximum nightly WST and median nightly RR, averaged across the preceding three-night window. Other parameters were excluded. Table 4 summarises the model performance metrics for the training and testing samples. Class 1 represented an 8-day long training instance extracted from day 10 to day 2 before SO. Class 0 represented a training instance extracted from all other 8-day long consecutive measurements. The training set consisted of 40 days of measurements from 66 participants with a 70:30 train-test split. Sensitivity is reflected in the recall of class 1, whereas specificity is determined by the recall of Class 0. Training the algorithm to detect COVID-19 1 day before SO did not improve recall (data not shown).

In the test set, the algorithm detected 68% of COVID-19 cases 2 days prior to SO.

DISCUSSION
Our main objective was to assess the use of existing medical-grade technology in the early detection of changes in physiological parameters related to COVID-19, thereby facilitating early isolation and testing of potentially affected individuals to limit the spread of the SARS-CoV-2 virus. Our RNN algorithm, trained and tested using a 70:30 split, identified 68% of COVID-19 cases...
up to 2 days before SO in 66 participants with an accurate false-positive rate and laboratory-confirmed cases of SARS-CoV-2. Therefore, we demonstrated that a wearable sensor bracelet implemented alongside a machine-learning model has the potential to detect COVID-19 infections prior to SO.

Our research is one of the first prospective cohort studies using wearable sensor technology to gather real-time continuous physiological data on which a machine-learning algorithm for COVID-19 presymptomatic detection was trained. Previous studies have evaluated the use of different wearable devices and machine learning

| Table 2  | Clinical characteristics of participants who contracted COVID-19 stratified according to whether they did (compliant group) or did not (non-compliant group) wear the bracelet regularly |
|----------------|---------------------------------------------------------------------------------|------------------|-----------------|-----------------|
| Variables (n) | Compliant group (n=66) | Non-compliant group (n=61) | Test statistic | Significance (p value) |
| Sex ratio (F:M) | 45:21 | 29:32 | $\chi^2(1)=4.74$ | 0.030 |
| Mean age, years (SD) | 42.88 (5.59) | 44.54 (5.60) | $F(1, 116)=2.85$ | 0.094 |
| BMI, kg/m² (SD) | 23.75 (3.69) | 25.81 (4.06) | $F(1, 116)=10.39$ | 0.002* |
| Hospitalisation rate | 3 | 7 | $\chi^2(1)=0.64$ | 0.425 |
| Smoking status, N (never: current: past smoker) | 57:4:5 | 36:6:7 | $\chi^2(2)=3.03$ | 0.22 |
| N of household contacts with COVID-19 | 35 | 18 | $\chi^2(1)=2.39$ | 0.123 |
| N of work colleagues with COVID-19 | 28 | 21 | $\chi^2(1)=0$ | 1 |
| COVID-19 symptoms: | | | | |
| Fever | 17 | 23 | $\chi^2(1)=0.89$ | 0.344 |
| Chills | 14 | 11 | $\chi^2(1)=0.62$ | 0.432 |
| Cough | 26 | 30 | $\chi^2(1)=0.25$ | 0.616 |
| Runny nose | 26 | 25 | $\chi^2(1)=0.01$ | 0.938 |
| Difficulty breathing | 11 | 10 | $\chi^2(1)=0.39$ | 0.530 |
| Loss of the sense of smell | 26 | 24 | $\chi^2(1)=0.37$ | 0.543 |
| Loss of the sense of taste | 20 | 22 | $\chi^2(1)=0.02$ | 0.896 |
| Chest pressure | 7 | 10 | $\chi^2(1)=0.22$ | 0.636 |
| Sore throat | 18 | 19 | $\chi^2(1)=0.00$ | 1 |
| Muscle pain | 27 | 32 | $\chi^2(1)=0.29$ | 0.593 |
| Headache | 44 | 29 | $\chi^2(1)=7.88$ | 0.005 |
| Fatigue | 27 | 38 | $\chi^2(1)=2.24$ | 0.135 |
| Malaise | 19 | 25 | $\chi^2(1)=0.18$ | 0.670 |
| Diarrhoea | 13 | 13 | $\chi^2(1)=0.02$ | 0.896 |
| Sickness | 9 | 5 | $\chi^2(1)=1.29$ | 0.256 |
| Vomiting | 1 | 5 | $\chi^2(1)=1.89$ | 0.169 |
| Hospitalisation | 3 | 7 | $\chi^2(1)=0.64$ | 0.425 |
| Long-term effects of COVID-19 (≥10 day) | 5 | 15 | $\chi^2(1)=5.69$ | 0.017 |
| Mean symptom duration | 8.54 (5.10) | 10.16 (10.98) | $F(1, 116)=1.31$ | 0.254 |

*Indicates p≤0.002, significant difference with Bonferroni correction. BMI, body mass index.
to identify COVID-19 infections based on self-reported COVID-19 infections.²⁵⁻³¹ Mishra et al.,⁹ for example, evaluated the use of resting HR data from 32 infected Fitbit users to detect COVID-19 cases in real time and identified 62.5% of the cases before SO. Similarly, Miller et al.²² used RR, HR and HRV data from 271 WHOOP strap wearers to identify 20% of participants who developed COVID-19 before SO and 80% by day 3 after SO.

Only laboratory-confirmed SARS-CoV-2 infections were used in this study to ensure more conclusive results. Our RNN algorithm detected 68% of laboratory-confirmed cases before SO, with additional statistical analyses revealing significant changes in the HR, HRV and WST, across the disease trajectory. Furthermore, our algorithm included more concurrent physiological parameters than previous studies, such as nightly RR, WST and cardiac data.⁷ ⁹ ³¹⁻³⁵ Unlike previous studies that performed retrospective measurements, our system could detect infections before SO. Uniquely, our research repurposed a previously existing CE-marked medical device for a novel purpose, illustrating a relatively inexpensive technique for detecting presymptomatic COVID-19. This

Table 3  Multi-level linear mixed models reveal the relationship between COVID-19 phases and physiological parameters

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Respiratory rate</th>
<th>Heart rate</th>
<th>Heart rate variability (SDNN)</th>
<th>Heart rate variability (RMSSD)</th>
<th>Heart rate variability ratio</th>
<th>Wrist skin temperature</th>
<th>Skin perfusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.10† (0.26)</td>
<td>55.43† (0.83)</td>
<td>59.64† (1.43)</td>
<td>43.71† (1.16)</td>
<td>0.50† (0.02)</td>
<td>35.32† (0.06)</td>
<td>−0.01† (0.00)</td>
</tr>
<tr>
<td>COVID-19 phase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>Reference group</td>
<td>Reference group</td>
<td>Reference group</td>
<td>Reference group</td>
<td>Reference group</td>
<td>Reference group</td>
<td>Reference group</td>
</tr>
<tr>
<td>Incubation</td>
<td>0.02 (0.06)</td>
<td>0.87† (0.29)</td>
<td>−1.48* (0.59)</td>
<td>−0.37 (0.48)</td>
<td>−0.01* (0.01)</td>
<td>0.13† (0.04)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Presymptomatic</td>
<td>0.14 (0.12)</td>
<td>1.00† (0.36)</td>
<td>−1.70* (0.64)</td>
<td>−0.75 (0.53)</td>
<td>−0.02* (0.01)</td>
<td>0.18† (0.05)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Symptomatic</td>
<td>1.00† (0.18)</td>
<td>2.15† (0.48)</td>
<td>−1.45* (0.73)</td>
<td>0.12 (0.51)</td>
<td>0.00 (0.01)</td>
<td>0.30† (0.05)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Recovery</td>
<td>0.10 (0.06)</td>
<td>0.87† (0.22)</td>
<td>−0.92 (0.51)</td>
<td>0.04 (0.44)</td>
<td>0.00 (0.01)</td>
<td>0.20† (0.03)</td>
<td>0.00 (0.00)</td>
</tr>
</tbody>
</table>

Unstandardised β-coefficient values reported, with SEs in brackets. *P<0.05. †0.007, respectively, with Bonferroni correction. RMSSD, root mean square of successive differences; SDNN, SD of the normal-to-normal interval.

Figure 5  The wearable device can detect changes in five physiological parameters across the clinical course of COVID-19. The values of each physiological parameter (with 95% CIs) collapsed across individuals (n=66) were normalised using baseline measurements and are shown centred around participant-reported symptom onset (SO). SDNN, SD of the normal-to-normal interval.
machine-learning algorithm can be applied to any sensor device that measures the same physiological parameters.

Our findings suggest that a wearable-informed machine-learning algorithm may serve as a promising tool for presymptomatic or asymptomatic detection of COVID-19. However, RT-PCR testing remains the most effective method to confirm COVID-19 infections. A systematic review of wearable sensors in detecting COVID-19 reported these investigations as promising but also highlighted the need for investigations in broader populations. Based on this interim analysis, a 20,000-person randomised controlled trial is underway to test the real-time efficacy of the RNN algorithm which can act on real-time machine-learning-driven alerts about the likelihood of a COVID-19 infection before symptoms are reported. The initial results from this larger trial are expected in December 2022, with a wider validation and more practical implications of the first presented data approach. In addition, detecting other illnesses using wearable-informed machine-learning algorithm is promising.

The strengths of our study include its population-based design and recruitment from a well-defined and well-characterised healthy cohort. A small subsample of COVID-19 positive users with sufficient high-quality data (wearing the Ava-bracelet ≥28 days prior to SO), reliance on data from a single national centre and lack of ethnic diversity may limit the generalisability of our findings. Additionally, we could not exclude imprecision or misclassification errors related to the symptoms experienced, dates of SO and/or SE. We acknowledge that our sensitivity was less than 80%. We expect to improve the algorithm’s performance further in a larger cohort within the setting of the COVID-RED study. Finally, one could argue that about half of the individuals identified as positive by the bracelet did not show SARS-CoV-2 infection in subsequent laboratory testing, and an unnecessary testing burden could arise from this fact. The positivity rates of PCR testing (ie, approximately 15%, depending on disease prevalence) in symptomatic outpatients routinely tested during the pandemic were considerably lower than the 50% observed in asymptomatic Ava-bracelet users. Hence, the Ava-bracelet could be regarded as progress when compared with the current testing routine.

Overall, the COVI-GAPP study showed that presymptomatic detection of COVID-19-related changes in physiological parameters using a sensor bracelet is feasible. We found significant changes in HR, HRV and WST occurring in COVID-19 positive patients during the presymptomatic period compared with baseline measurements, over and above the effects of intrapersonal variability. A novel machine-learning algorithm detected 68% of laboratory-confirmed SARS-CoV-2 infections 2 days before SO. Wearable sensor technology is an easy-to-use, low-cost method for enabling individuals to track their health and well-being during a pandemic. Our research shows how these devices, partnered with artificial intelligence, can push the boundaries of personalised medicine and detect illnesses prior to SO, potentially reducing virus transmission in communities. Future research should focus on how medical-grade wearable sensor technology can aid in combating the current pandemic by monitoring sensor data.

**Author affiliations**

1. Dr Risch Medical Laboratory, Vaduz, Liechtenstein
2. Central Laboratory, Canton Hospital Graubünden, Chur, Switzerland
3. Ava AG, Zurich, Switzerland
4. Faculty of Medical Sciences, Private University in the Principality of Liechtenstein, Triesen, Liechtenstein
5. Cardiovascular Research Institute Basel (CRIB), University Hospital Basel, University of Basel, Basel, Switzerland
6. Department of Metabolism, Digestive Diseases and Reproduction, Imperial College London, London, UK
7. Institute of Laboratory Medicine and Pathobiochemistry, Molecular Diagnostics, Philipps University Marburg, Marburg, Germany
8. Department of Cardiology and University Center of Cardiovascular Science, University Heart and Vascular Center Hamburg, Hamburg, Germany
9. AvA AG, Zurich, Switzerland
10. Department of Psychology, University of Fribourg, Fribourg, Switzerland
11. Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland
12. Julius Clinical, Zeist, The Netherlands
13. UMC Utrecht, Utrecht, The Netherlands
14. Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands
15. Takeda Pharmaceuticals, Digital Clinical Devices, Cambridge, UK
16. Roche Diagnostics Nederland B.V, Almere, The Netherlands
17. Population Health Research Institute, McMaster University, Hamilton, Ontario, Canada
18. Center of Laboratory Medicine, University Institute of Clinical Chemistry, University of Bern, Bern, Switzerland

**Acknowledgements** We thank the GAPP participants who enrolled in this study. Additionally, the authors thank the following for their contributions to the study: The local study team in Vaduz, FL, the different teams at the Dr Risch Medical Laboratories in Vaduz and Buchs, CH. We would also like to thank the Coobx AG.

**Table 4** Performance metrics of the algorithm in the detection of COVID-19 2 days prior to symptom onset class 0 represented an 8-day long training instance extracted from day 10 to day 2 before SO.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>0</td>
<td>0.60</td>
<td>0.45</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.60</td>
<td>0.73</td>
<td>0.66</td>
</tr>
<tr>
<td>Test set</td>
<td>0</td>
<td>0.50</td>
<td>0.36</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.54</td>
<td>0.68</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Class 0 represented a training instance extracted from all other 8 days long consecutive measurements (eg, SO-11 to SO-3). The training set consisted of 40 days measurements from 66 participants with 70:30 train-test split. Sensitivity is reflected in the recall of class 1, while specificity is determined by the recall of class 0. SO, symptom onset.
Finally, special thanks are due to the Princely House of Liechtenstein, which gave

**Contributors**

All authors critically reviewed and approved the final version of this manuscript and had final responsibility for the decision to submit for publication. Conceptualisation: MRi, HR, RT, PK, TBB, BF, MM, GSD, AD, SM, MC, DEG and DC; BMG; TL; Funding: KG, NW, CR, DH, MRi, VK and DL; Visualisation: KG and FP; Writing—original draft: MRo, KG, FP, MC, DC, BMG; LR; Writing—review and editing: MRi, KG, SA, OCW, MW, CR, DH and TL; Supervision: MRi, SA, HR, RT, MC, DEG, DC, BMG; LR; Validation: KG, NW, CR, DL; Visualisation: KG and FP; Writing—original draft: MRo, KG, FP, MC, DC, BMG; LR; Writing—review and editing: MRi, KG, SA, OCW, MW, CR, DH and TL; Supervision: MRi, SA, HR, RT, MC, DEG, DC, BMG; LR; Funding: The COVI-GAPP study received grants from the Innovative Medicines Initiative (IMI grant agreement number 101005177), the Princely House of Liechtenstein, the government of the Principality of Liechtenstein, and the Hanela Foundation in Aarau (Switzerland). None of the funders played a role in the study design, data collection, data analysis, data interpretation, writing of the report, or decision to publish.

**Competing interests**

LR, and MR are key shareholders of the Dr Risch Medical Laboratory. DC has received consulting fees from Roche Diagnostics, outside of the current work. MR, DL, YK, AM, MC, and BMG are employed by AvA AG. The other authors have no financial or personal conflicts of interest to declare.

**Patient and public involvement**

Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication**

Consent obtained directly from patient(s)

**Ethics approval**

This study involves human participants and was approved by KEK, Zurich, Switzerland (BASEC 2020-00786).

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**Data availability statement**

Data are available on reasonable request. Anonymised data that underlie the results reported in this article are available upon justifiable request to the corresponding author.

**Supplemental material**

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to localised regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access**

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits non-commercial reproduction and distribution, provided the author(s) and source are credited.
permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc-4.0/.

ORCID iDs
Kirsten Grossmann http://orcid.org/0000-0002-0460-3736
Ornela C Weideli http://orcid.org/0000-0003-2339-9440
Harald Renz http://orcid.org/0000-0003-0602-7215
Raphael Tverenbold http://orcid.org/0000-0003-3814-6542
Martina Rothenbühler http://orcid.org/0000-0001-8009-4104
Andjela Markovic http://orcid.org/0000-0003-2104-9543
Marianna Mitratza http://orcid.org/0000-0002-1573-2015
George S Downward http://orcid.org/0000-0001-6642-7048
Ariel Dowling http://orcid.org/0000-0002-7889-4978
Diederick E Grobbbee http://orcid.org/0000-0003-4472-4468
Lorenz Risch http://orcid.org/0000-0003-2692-6699

REFERENCES
Supplementary Materials

Supplement to: “Investigation of the use of a sensor bracelet for the pre-symptomatic detection of COVID-19: An interim analysis of a national cohort study (COVI-GAPP)”.
Supplementary Material and Methods

Our primary aim was to understand how the coronavirus disease 2019 (COVID-19) affects physiological parameters measured by a wearable device and, subsequently, whether these parameter changes could help in detecting a pre-symptomatic infection. In particular, we investigated how heart rate (HR), respiratory rate (RR), heart rate variability (HRV), wrist-skin temperature (WST), and skin perfusion deviated from baseline measurements during four infection-related periods: the incubation period, the pre-symptomatic period, symptomatic infection period, and the recovery period. We categorized daily parameter measurements as occurring in the baseline period if the day ($d$) was more than 10 days prior to symptom onset (SO; i.e., $d > SO - 10$). Relatedly, we defined the incubation period as $SO - 10 \leq d < SO - 2$ and the pre-symptomatic period as $SO - 2 \leq d < SO$. Because participants’ reported symptom duration varied, measurements fell into the symptomatic infection category if $SO \leq d \leq SE$. Finally, parameters collected after symptom end (SE) were classified as in the recovery period (i.e., $d > SE$).

The Wearable Device and Physiological Parameter Specification

The Ava Fertility Tracker (version 2.0; Ava AG, Switzerland) is an United States Food and Drug Administration (FDA) cleared and conformité européenne (CE) certified fertility aid bracelet that complies with international regulatory requirements and applicable standards. The wrist-worn tracker consists of three sensors: a temperature sensor; an accelerometer; and a photoplethysmograph (PPG). The Ava-bracelet saves data every 10 seconds and requires at least four hours of relatively uninterrupted sleep to record enough data for pre-processing and analysis. Upon waking, the user taps a button in the complementary smartphone app to initiate the previous night’s raw data transfer from the Ava-bracelet to the system’s backend database via Bluetooth Low Energy (BLE). The data then undergoes pre-processing according to proprietary manufacturer algorithms to remove potential artifacts, detect the user’s sleep stages, and identify nightly physiological parameters. In addition to the algorithm-derived fertility indication, the post-processing values for HR, WST, RR, sleep quantity, sleep quality, and HRV ratio are then sent back to the complementary app and displayed to the user. The device’s sensors responsible for recording the raw data are described in detail below as well as show in Figure S1.

Built into the Ava-bracelet’s internal hardware, the accelerometer detects and records the wearer’s movement in three-dimensional space. A proprietary machine learning algorithm ingests nightly movement data to determine sleep stages. In addition to reporting the user’s duration of sleep in-app, it also assigns her a nightly sleep quality score consisting of the percentage of combined deep and Rapid Eye Movement (REM) sleep. Although other researchers have examined COVID-19’s impact on sleep using wearable devices with mixed or inconclusive results, since sleep quality and quantity were not among our pre-defined primary objectives we did not analyse results from the accelerometer data.

A temperature sensor constitutes the Ava-bracelet second sensor and provided data for evaluating COVID-19 related changes in wrist skin temperature (WST). Despite the device reading temperature at a distal point compared to core body temperature, recent research has demonstrated the Ava-bracelet’s ability to continuously measure temperature throughout the night results in more sensitive readings than oral point estimates and enables its machine learning algorithms to detect more ovulation-related changes in temperature. These findings suggest the medical grade device’s ability to sense fluctuations in WST related to an infection would similarly benefit from its repeated sampling over the course of sleep and may outperform an oral or forehead reading taken only once at point of care (POC). Limited evidence conducted early on during the COVID-19 pandemic attests to WST’s potential superior usage in detecting infection-based fluctuations; WST for 528 patients read by a noncontact infrared thermometer proved more stable and less prone to environmental factors (e.g., walking or bicycling to POC) than tympanic and forehead measurements in some contexts. Thus, given prior research on the Ava-bracelet’s measurement accuracy compared to oral temperature and on WST’s importance in triaging COVID-19 patients, we relied on the device’s temperature sensor to provide nightly WST readings for analysing how temperature changes across a symptomatic SARS-CoV-2 infection.

A PPG comprises the Ava-bracelet’s final sensor. The PPG sensor employs a light emitting diode (LED) current to send infrared light through the user’s skin to detect inter-beat intervals (IBIs). The light reflects off or is absorbed by the blood; how much light bounces back to the sensor can signal the wearer’s current cardiac rhythms. Based on the time cadence for variance in the reflected light, proprietary algorithms can determine the user’s HR, RR,
perfusion and IBI; in turn, the IBI can inform calculations for various metrics of HRV. While HR consists of the number of heart beats per minute, HRV describes the fluctuation in time intervals between consecutive heartbeats.\(^{10}\) It can vary in both frequency- and time-domains, resulting in more than 20 possible metrics for quantifying the heart’s activity.\(^{10}\) Since examining all HRV metrics would have proven practically and statistically infeasible, we focused on two time- and one frequency-domain measurements. The first time-domain measure of HRV, the standard deviation of the NN interval (SDNN), quantifies sympathetic and parasympathetic nervous system activity in ms; it describes how much variability exists in the interval between normal sinus beats.\(^{10}\) A lower SDNN corresponds to impaired cardiac health\(^{10}\), with recent research offering conflicting evidence about SDNN’s changes in COVID-19 patients. While some studies demonstrated an increase in SDNN among COVID-19 patients\(^{11}\), others have found changes in SDNN dependent upon disease severity.\(^{12}\) Regardless of the effect’s direction, we expected an individual suffering from COVID-19 would exhibit deviations from their baseline SDNN during an active infection and included it in our analyses. A second time-domain measurement of HRV, the root mean square of successive differences (RMSSD), examines the variability between normal heartbeats. Increased RMSSD has previously been shown to be associated with severe infection, including septic shock and COVID-19.\(^{11,13}\) Thus, we focused on RMSSD changes across the incubation, pre-symptomatic, symptomatic and recovery phases compared to differences included it in our analyses. A second time-domain measurement of HRV, the root mean square of successive differences (RMSSD), examines the variability between normal heartbeats. Increased RMSSD has previously been shown to be associated with severe infection, including septic shock and COVID-19.\(^{11,13}\) Thus, we focused on RMSSD changes across the incubation, pre-symptomatic, symptomatic and recovery phases compared to participants’ baseline measurements in our analysis. The final HRV parameter we examined, the HRV ratio, constitutes a frequency-domain measurement; it indicates the ratio of HR oscillations in the low-frequency (LF; 0-0.04 Hz) to those in the high-frequency (HF; 0.15-0.4 Hz) bands\(^{10,14}\). Patients with severe COVID-19 infection have exhibited a higher HRV ratio than mildly infected participants\(^{12}\); leading us to examine this physiological parameter in our analyses.

**Data Processing and Multi-level Model Specification**

We performed all data processing and analysis using R (R Core Team, v3.6.1\(^{15}\)) and Python (Python Software Foundation, v3.6\(^{15}\)). In keeping with data cleaning practices described by the manufacturer in previous publications,\(^3\) we excluded the first 90 and the last 30 minutes of data from each night a priori from our analysis; transitions from waking to sleeping and vice versa can result in greater variation in physiological parameters measured by the Ava-bracelet, thereby leading to less stable readings. To further reduce artificial fluctuations in the data due to potential measurement error and consistent with best practices\(^{16}\), each physiological parameter underwent locally estimated scatterplot smoothing (LOESS) prior to analysis.

Next, we ran a series of multi-level models with random intercepts and random slopes to determine differences in physiological parameters during the infection-related periods compared to baseline, accounting for the nesting of repeated measurements during an infection period and within an individual. Given our continuous criterion, we used the “lme” function with residual maximum likelihood estimation (REML) and Satterthwaite degrees of freedom in the open-source R packages “lme4”\(^{18}\), “lmerTest”\(^{19}\), and “optimx”\(^{20}\) to model our outcomes of interest. Four dummy-coded variables were created, indicating to which infection period a given measurement belonged (1= Belonging to that Period, 0=Not belonging to that period). The reference baseline period measurements were encoded as 0 across all four dummy variables. Our reported results include the unstandardized regression coefficients for each effect. When multiple models were possible for the same parameter, we chose the model using the percentile of data (stable maxima) with the best fit; we determined best fit by comparing the two models using an analysis of variance (ANOVA) test and selecting the model with the significantly lower Akaike Information Criterion (AIC). In instances where the models were not significantly different from each other, we chose the model that included more data (e.g., the 99% percentile of data versus the 90th percentile).

In an effort to provide some context for the magnitude of our significant effects, we report the intraclass correlation coefficient (ICC) for each of the null models associated with changes in physiological parameters over the course of a COVID-19 infection. The ICC indicates how much variance in an outcome occurs due to between group differences\(^{21-23}\); in the context of the current study, the ICC presents a picture of how a given physiological parameter varies due to participant-level characteristics versus the within-subject course of a COVID-19 infection.

To ensure a family-wise alpha level less than or equal to .05, we implemented a Bonferroni correction for the seven total parameters we analyzed and evaluated effect significance using this new level of p=.007. We adjusted how we defined marginal significance accordingly (i.e., .007≤p≤.05). We used the Bonferroni-corrected significance level throughout the paper.
Supplementary Results

The ICCs and random effects variance estimates for each of the seven multi-level models can be found in Table S1. In brief, most physiological parameters had high levels of variance which could be attributed to between participant differences rather than within subject changes due to COVID-19 infection.

For most physiological parameters, observed variance in the outcome resulted largely from a participant’s own stability in readings over time. All cardiac parameters showed similar ICCs, ranging from 0.71 (RMSSD) to 0.77 (SDNN); this means that, depending on the parameter, 71-77% of the variance in outcome was due to between participant differences. Regardless of infection phase, a given participant’s nightly cardiac measurements were more similar to one another than random chance. RR showed an even higher ICC; 88% of all observed variance in RR was attributable to between participant differences. A maximum of 22% of variance could be due to within participant changes. The multi-level model testing the effect of infection phase on nightly RR reveals only a significant difference between the symptomatic period and baseline (see Table 3); all other phases do not differ significantly from baseline, illustrating the lack of overall variability due to a COVID-19 infection and emphasizing RR’s stability over time within an individual participant.

On the other end of the spectrum, only wrist skin temperature and perfusion had low ICC’s (0.01 and 0.05, respectively); said differently, a given participant’s perfusion or temperature measurements over time were not more similar to each other than would be expected from a random selection of that same parameter across all participants. As perfusion did not show phase-based changes in COVID-19 infection (see Table 3), it may be that another unaccounted for factor contributes to outcome measurements. Neither the participant’s own repeated measurements nor the disease trajectory appear to significantly influence a given night’s perfusion data. In contrast, since wrist skin temperature significantly differed from baseline across all other phases of a COVID-19 infection (see Table 3), it appears that the disease itself contributes more to a given night’s temperature readings than the stability in a participant’s own repeated measurements; almost all of the observed variance in nightly skin temperature occurs due to within participant differences (e.g., changes in their physiology over the course of the infection). Examining ICC values for each physiological parameter of interest provides greater context into the relative effect of potential phase-based changes in outcome variables as well as the residual variance attributable to the participant themselves.

Supplementary Tables and Figures

Supplementary Table 1. Intraclass correlation coefficients (ICCs) calculated based on the variance estimates for random effects of the null models predicting each of the seven physiological parameters of interest.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Between Participant Variance (SD)</th>
<th>Variance of the Residuals (SD)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist Skin Temperature</td>
<td>0.34 (0.59)</td>
<td>35.65 (5.97)</td>
<td>0.01</td>
</tr>
<tr>
<td>Heart Rate</td>
<td>43.59 (6.00)</td>
<td>13.53 (3.68)</td>
<td>0.76</td>
</tr>
<tr>
<td>Heart Rate Variability (SDNN)</td>
<td>121.64 (11.03)</td>
<td>36.08 (6.08)</td>
<td>0.77</td>
</tr>
<tr>
<td>Heart Rate Variability (RMSSD)</td>
<td>82.98 (9.06)</td>
<td>33.79 (5.81)</td>
<td>0.71</td>
</tr>
<tr>
<td>Heart Rate Variability Ratio</td>
<td>1.16 (1.08)</td>
<td>0.40 (0.63)</td>
<td>0.74</td>
</tr>
<tr>
<td>Respiratory Rate</td>
<td>4.48 (2.12)</td>
<td>0.64 (0.80)</td>
<td>0.88</td>
</tr>
<tr>
<td>Skin perfusion</td>
<td>3.8 e-05 (9.01)</td>
<td>6.75 e-04 (4.03)</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Supplementary Figure 1. The Ava Fertility Tracker contains three sensors (temperature, accelerometer and photoplethysmograph) that measure wrist skin temperature, heart rate, respiratory rate, heart rate variability and skin perfusion simultaneously.

Study protocol

The study protocol can be downloaded here.

Supplementary references


