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≤d<SO-2 and the pre-symptomatic period as SO-2≤d<SO. Because participants' reported symptom duration varied, measurements fell into the symptomatic infection category if SO≤d≤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.[1,2] The wrist-worn tracker consists of three sensors: a temperature sensor; an accelerometer; and a photoplethysmograph (PPG).[3] 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[4–7], 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.[8] 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.[9] 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 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; 0·04-0·15 Hertz [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[16]). 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[17], 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.

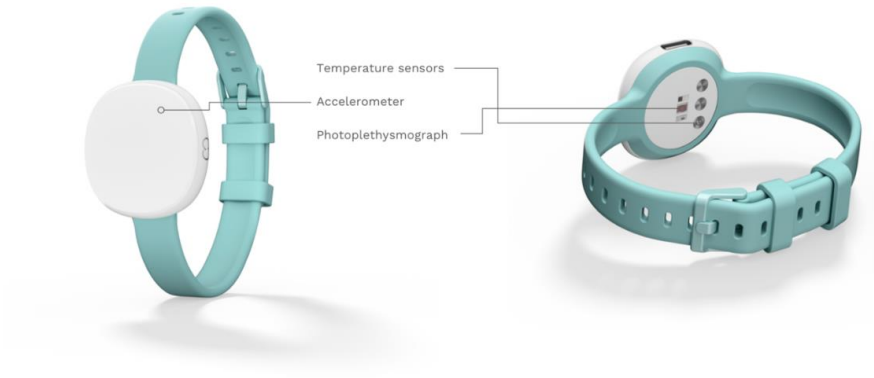
On the other end of the spectrum, only wrist skin temperature and perfusion had low ICC's (0·01 and 0·05, respectively); said differently, a given participant's perfusion or temperature measurements over time were not more similar to each other than would be expected from a random selection of that same 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**.**

Predictors	Between Participant Variance (SD)	Variance of the Residuals (SD)	ICC
Wrist Skin Temperature	0·34 (0·59)	35·65 (5·97)	0·01
Heart Rate	43·59 (6·60)	13·53 (3·68)	0·76
Heart Rate Variability (SDNN)	121·64 (11·03)	36·08 (6·08)	0·77
Heart Rate Variability (RMSSD)	82·08 (9·06)	33·79 (5·81)	0·71
Heart Rate Variability Ratio	1·16 (1·08)	0·40 (0·63)	0·74
Respiratory Rate	4·48 (2·12)	0·64 (0·80)	0·88
Skin perfusion	3·8 e-05 (0·01)	6·75 e-04 (0·03)	0·05

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

1. French-Mowat E, Burnett J. How are medical devices regulated in the European Union? *J R Soc Med*. 2012;105 Suppl:22-28. doi:10.1258/jrsm.2012.120036
2. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *Lancet Digit Heal*. 2021;3(3):e195-e203. doi:10.1016/S2589-7500(20)30292-2
3. Goodale BM, Shilaih M, Falco L, Dammeier F, Hamvas G, Leeners B. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: Observational study. *J Med Internet Res*. 2019;21(4). doi:10.2196/13404
4. Quer G, Radin JM, Gadaleta M, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat Med*. 2021;27(1):73-77. doi:10.1038/s41591-020-1123-x
5. Mishra T, Wang M, Metwally AA, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat Biomed Eng*. 2020;4(12):1208-1220. doi:10.1038/s41551-020-00640-6
6. Bogu GK, Snyder MP. Deep learning-based detection of COVID-19 using wearables data. *medRxiv*. Published online 2021:2021.01.08.21249474. http://medrxiv.org/content/early/2021/01/09/2021.01.08.21249474.abstract
7. Shapiro A, Marinsek N, Clay I, et al. Characterizing COVID-19 and Influenza Illnesses in the Real World via Person-Generated Health Data. *Patterns*. 2021;2(1):100188. doi:10.1016/j.patter.2020.100188
8. Zhu TY, Rothenbühler M, Hamvas G, et al. The accuracy of wrist skin temperature in detecting ovulation compared to basal body temperature: Prospective comparative diagnostic accuracy study. *J Med Internet Res*. 2021;23(6):1-12. doi:10.2196/20710
9. Moraes JL, Rocha MX, Vasconcelos GG, Vasconcelos Filho JE, de Albuquerque VHC, Alexandria AR. Advances in photopletysmography signal analysis for biomedical applications. *Sensors (Switzerland)*. 2018;18(6):1-26. doi:10.3390/s18061894
10. Shaffer F, Ginsberg JP. An Overview of Heart Rate Variability Metrics and Norms. *Front Public Heal*.

2017;5(September):1-17. doi:10.3389/fpubh.2017.00258
11. Kaliyaperumal D, Rk K, Alagesan M, Ramalingam S. Characterization of cardiac autonomic function in COVID-19 using heart rate variability: A hospital based preliminary observational study. *J Basic Clin Physiol Pharmacol*. 2021;32(3):247-253. doi:10.1515/jbcpp-2020-0378
12. Pan Y, Yu Z, Yuan Y, et al. Alteration of Autonomic Nervous System Is Associated With Severity and Outcomes in Patients With COVID-19. *Front Physiol*. 2021;12(May):1-10. doi:10.3389/fphys.2021.630038
13. Chen WL, Kuo CD. Characteristics of Heart Rate Variability Can Predict Impending Septic Shock in Emergency Department Patients with Sepsis. *Acad Emerg Med*. 2007;14(5):392-397. doi:10.1197/j.aem.2006.12.015
14. Yilmaz M, Kayancicek H, Cekici Y. Heart rate variability: Highlights from hidden signals. *J Integr Cardiol*. 2018;4(5):1-8. doi:10.15761/jic.1000258
15. R Core Team (2019). language and environment for statistical computing. Published online 2019. https://www.r-project.org/
16. Python Software Foundation. http://www.python.org
17. Colantuoni C, Henry G, Zeger S, Pevsner J. Local mean normalization of microarray element signal intensities across an array surface: Quality control and correction of spatially systematic artifacts. *Biotechniques*. 2002;32(6):1316-1320. doi:10.2144/02326mt02
18. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1). doi:10.18637/jss.v067.i01
19. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models . *J Stat Softw*. 2017;82(13). doi:10.18637/jss.v082.i13
20. Nash JC, Varadhan R. Unifying optimization algorithms to aid software system users: Optimx for R. *J Stat Softw*. 2011;43(9):1-14. doi:10.18637/jss.v043.i09
21. Hox J. *Multilevel Analysis: Techniques and Applications. Quantitative Methodology Series.* New York, NY: Routledge.; 2002.
22. LaHuis DM, Hartman MJ, Hakoyama S, Clark PC. Explained Variance Measures for Multilevel Models. *Organ Res Methods*. 2014;17(4):433-451. doi:10.1177/1094428114541701
23. Raudenbush, S. W., & Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods: Applications and Data Analysis Methods (2nd Ed., Vol. 1).* Thousand Oaks, CA: SAGE; 2002.