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,1,4 Stefanie Aeschbacher,5 Ornella C Weideli,1 Marc Kovac,3 Fiona Pereira,6 Nadia Wohlwend,3 Corina Risch,3 Dorothea Hillmann,3 Thomas Lung,3 Harald Renz,7 Raphael Twerenbold,5,8 Martina Rothenbühler,6,9 Daniel Leibovitz,9 Vladimir Kovacevic,9 Andjela Markovic,4,9,10,11 Paul Klaver,12 Timo B Brakenhoff,12 Billy Franks,12 Marianna Mitratza13,14 George S Downward13,14 Ariel Dowling15 Santiago Montes,16 Diederick E Grobbee,13 Maureen Cronin,1 David Conen,17 Brianna M Goodale,9 Lorenz Risch2,3,18 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 ISRCTN51255782; 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. 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 presymptomatic 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 presymptomatic 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 analyze COVID-19-related changes in various physiological parameters across four infection-related periods: incubation, presymptomatic, 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 presymptomatic 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 to detect COVID-19 infection.

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.20 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).20–22 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 \( d > 10 \) days prior to SO (ie, \( d > SO-10 \)), the incubation period as \( SO-10 \leq d < SO-2 \), and the presymptomatic period as \( SO-2 \leq d < SO \). We chose a cut-off of −2 days based on previous reports of infected participants becoming contagious 2 days before SO.23 Because the participants’ reported symptom durations varied, the measurements were categorised into the symptomatic infection category if \( SO \leq d \leq SE \). 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.24 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 practices (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 output 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 (ie, $0.007 \leq p \leq 0.05$).

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 (eg, 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 (eg, 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
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 [3 11 18 19 25 26] 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:

\[
\text{overall\_recall} = \left( (\text{recall\_class\_1\_train} + \text{recall\_class\_0\_train}) \times 0.7 + (\text{recall\_class\_1\_test} + \text{recall\_class\_0\_test}) \times 0.3 \right) / 2
\]

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 phase of COVID-19.

**Patient and public involvement**
No patient or public involvement.

**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 \( p \geq 0.30 \)). 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=5.6) and most had never smoked (n=57; 86%). Their COVID-19 symptoms lasted for an average of 8.5 days (SD=5.6; 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 ($\beta_{\text{intercept}} = 15.1$ 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
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 (βpresymptomatic = −1.7 ms, SE=0.64; 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=0.01; p=0.0361) and presymptomatic periods (βpresymptomatic = −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.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total n=1163</th>
<th>COVID-19 n=127</th>
<th>No COVID-19 n=1036</th>
<th>Test statistic</th>
<th>Significance (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex ratio (F:M)</td>
<td>667:494</td>
<td>74:53</td>
<td>594:441</td>
<td>χ²(4)=0.40</td>
<td>0.982</td>
</tr>
<tr>
<td>Mean age, years (SD)</td>
<td>44.08 (5.57)</td>
<td>43.66 (5.64)</td>
<td>44.14 (5.56)</td>
<td>F(1, 1071)=0.59</td>
<td>0.444</td>
</tr>
<tr>
<td>BMI, kg/m² (SD)</td>
<td>24.72 (3.97)</td>
<td>24.74 (4.00)</td>
<td>24.72 (3.97)</td>
<td>F(1, 1071)=0.02</td>
<td>0.90</td>
</tr>
<tr>
<td>Smoking status, N (never: current: past smoker)</td>
<td>654:110:102</td>
<td>93:10:12</td>
<td>561:100:90</td>
<td>χ²(2)=2.38</td>
<td>0.304</td>
</tr>
<tr>
<td>N of household contacts with COVID-19</td>
<td>111</td>
<td>53</td>
<td>58</td>
<td>χ²(1)=127.94</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>N of work colleagues with COVID-19</td>
<td>279</td>
<td>49</td>
<td>230</td>
<td>χ²(3)=27.3</td>
<td>&lt;0.0001*</td>
</tr>
</tbody>
</table>

*Indicates p<0.002, significant difference with Bonferroni correction.
BMI, body mass index.

**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 2 days prior to SO.
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...
to identify COVID-19 infections based on self-reported COVID-19 infections. Mishra et al. 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. 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. 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.
‡P<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 which 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 combatting the current pandemic by monitoring sensor data.

**Table 4** Performance metrics of the algorithm in the detection of COVID-19 2 days prior to symptom onset class 1 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.

**Author affiliations**

1Dr Risch Medical Laboratory, Vaduz, Liechtenstein
2Central Laboratory, Canton Hospital Graubünden, Chur, Switzerland
3Dr Risch Medical Laboratory, Buchs, Switzerland
4Faculty of Medical Sciences, Private University in the Principality of Liechtenstein, Triesen, Liechtenstein
5Cardiovascular Research Institute Basel (CRIB), University Hospital Basel, University of Basel, Basel, Switzerland
6Department of Metabolism, Digestive Diseases and Reproduction, Imperial College London, London, UK
7Institute of Laboratory Medicine and Pathobiology, Molecular Diagnostics, Philipps University Marburg, Marburg, Germany
8Department of Cardiology and University Center of Cardiovascular Science, University Heart and Vascular Center Hamburg, Hamburg, Germany
9Ava AG, Zurich, Switzerland
10Department of Psychology, University of Fribourg, Fribourg, Switzerland
11Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland
12Julius Clinical, Zeist, The Netherlands
13UMC Utrecht, Utrecht, The Netherlands
14Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands
15Takeda Pharmaceuticals, Digital Clinical Devices, Cambridge, UK
16Roche Diagnostics Nederland B.V, Almere, The Netherlands
17Population Health Research Institute, McMaster University, Hamilton, Ontario, Canada
18Center of Laboratory Medicine, University 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...
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: MR, HR, RT, PK, TBB, BF, MM, GD, SM, DE, DC and BM; LR; Data curation: KG, SA, MR and BM; Formal analysis: MRo, KG, MRo, DL, VK and BM; Funding acquisition: MRo, MC, DC and LR; Investigation: KG, SA, OCW and MM; Methodology: MRo, PK, TBB, GD, SM, MG, DC, DE, BMG and LR; Project administration: KG, OCW, SM, MC and DE; Resources: MRo, KG, SA, OCW, MK, CR, DH and TL; Supervision: MR, SA, HR, RT, MG, DE, DC, BMG and LR; Validation: KG, NW, CR, DH and DL; Visualization: KG and FP; Writing—original draft: MRo, KG, DP, MC, DC, BMG and LR; Writing—review and editing: MR, KG, SA, OCW, MK, CR, DH and TL; HR, RT, MR, VK, AM, DL, PK, TBB, BF, MM, GD, SM, AM, DC, DE, BMG and LR. All authors read and agreed to the publisher version of the manuscript. LR is acting as guarantor.

**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 Hansa 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, MR and MK are key shareholders of the Dr Risch Medical Laboratory. DC has received consulting fees from Roche Diagnostics, outside of the current work. MR, DL, VK, 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, or 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 on 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 local 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 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
Ornella C Weideli http://orcid.org/0000-0003-2389-9440
Harald Renz http://orcid.org/0000-0003-6062-7215
Raphael Twerenbold http://orcid.org/0000-0003-2692-6699
Harald Renz http://orcid.org/0000-0003-0602-7215

5 Centers for Disease Control and Prevention. Similarities and difference between flu and COVID-19 [Accessed March 2021].
29 Nestor B. Dear Watch should I get a COVID-19 test? designing deployable machine learning for wearables, 2021. https://www. medrxiv.org/content/10.1101/2021.05.11.21257052v1

See: http://creativecommons.org/licenses/by-nc/4.0/.