PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Investigation of the use of a sensor bracelet for the pre-symptomatic detection of changes in physiological parameters related to COVID-19: an interim analysis of a prospective cohort study (COVI-GAPP)
AUTHORS	Risch, Martin; Grossmann, Kirsten; Aeschbacher, Stefanie; Weideli, Ornella; Kovac, Marc; Pereira, Fiona; Wohlwend, Nadia; Risch, Corina; Hillmann, Dorothea; Lung, Thomas; Renz, Harald; Twerenbold, Raphael; Rothenbuehler, Martina; Leibovitz, Daniel; Kovacevic, Vladimir; Markovic, Andjela; Klaver, Paul; Brakenhoff, Timo; Franks, Billy; Mitratza, Marianna; Downward, George; Dowling, Ariel; Montes, Santiago; Grobbee, Diederick; Cronin, Maureen; Conen, David; Goodale, Brianna; Risch, Lorenz; Risch, Lorenz

VERSION 1 – REVIEW

REVIEWER	Chan, P Eastern Health, Melbourne, ICU
REVIEW RETURNED	09-Nov-2021

GENERAL COMMENTS	Well presented. No issues

REVIEWER	Li, Xiao Case Western Reserve University
REVIEW RETURNED	30-Nov-2021

GENERAL COMMENTS	With 127 confirmed COVID-19 cases out of 1163 participants recruited in this study, the authors of this manuscript investigated the application of using wearable devices (the Ave bracelet) in detecting COVID-related physiological changes. They further developed an RNN-based machine learning approach for pre-symptomatic COVID-19 detection and showed that their algorithm could identify 68% of COVID-19 positive participants two days before symptom onset. The strengths of this study include the large sample size and the fact that all the COVID-19 cases have been laboratory confirmed. In addition, a follow-up study is currently underway, which will potentially validate this algorithm and further strength this observation. I have several suggestions regarding the revision of this manuscript which are described below. 1. Please provide more details regarding the data and data process. As the author mentioned, the bracelet saves data every 10s and requires at least four hours of relatively uninterrupted sleep. In which resolution was each parameter been collected and analyzed? Are there missing data in the data collection? If yes,

	what was the distribution, and how did the analysis handle the missing data? 2. The authors stated that the RNN-based approach detected 68% of COVID-19 cases up to two days before SO in 66 participants with an accurate false-positive rate and laboratory-confirmed cases of SARS-CoV-2. It would be nice to show whether there are any features associated with the prediction success? For example, do the asymptomatic patients (3 patients) behave differently from those with severe symptoms or were hospitalized (11 patients)? It is probably a good idea to include more details in the model specification and algorithm performance section in the results section. 3. According to Table 4, it seems that the sensitivity of detecting COVID was evaluated by the recall of class 1 (i.e., day 10 – day 2 before SO). What was the rationale for determining this period? What if it includes the period between day 2 to SO? Is it going to further improve the sensitivity? Minor points: 1. Probably a good idea to have consistent terms in the paper. For example, 'respiration rate' vs. 'breathing rate' were both used in the article. 2. Line #26, how often is the follow-up for the study? 3. Line 45, 'SE' used before the first-time introduction (line 46)

REVIEWER	Channa, Asma University POLITEHNICA of Bucharest, Computer Science
REVIEW RETURNED	28-Dec-2021

GENERAL COMMENTS	The paper proposed a methodology to detect COVID-19 patients from the data of wearable sensors (bracelet). The authors used some existing well-known machine learning algorithms to detect COVID-19 patients. Adding very small recommendations for the authors - 1. References for standardized algorithms to be added in the manuscript – SVM, k-NN etc. For instance, previously other ML algorithms have been used to detect COVID symptoms that must be included in related work. 2. A brief write up must be included that discuss the reasons behind the accuracy being less than 80% 3. Future work and limitations should be extended.

REVIEWER	Mazumder, Nirmal Manipal Academy of Higher Education
REVIEW RETURNED	14-Jan-2022

GENERAL COMMENTS	The manuscript is well written and very important study at present situation. The manuscript can be accepted after minor revision. 1. The device is well recognized and used. Did the author use the device as it is or any hardware improvement made ? 2. The study is a kind of prediction model. How reliable it is ? Is it easy to use for common people? 3. what is the sensitivity and selectivity in covid detection ? 4. currently RT PCR based detection is suggested. can it be replaced by the proposed analysis ?

REVIEWER	Vasireddy, Deepa Pediatric Group of Acadiana, Pediatrics
REVIEW RETURNED	02-Feb-2022

GENERAL COMMENTS	Kindly address the following: 1. Why was antibody testing chosen over RT-PCR 2. Can the authors comment on generalisability of these findings if other similar devices were used. 3.After identifying pre symptomatic individuals through such devices, are the authors trying to convey that the patients could quarantine or get tested early on as the ultimate possible utility of the study. 4.What are the economic costs to the patient to own such a device. False positive pick up rate and possible unnecssary testing burden with such cases?

VERSION 1 – AUTHOR RESPONSE

Reviewer #1 Dr. P Chan, Eastern Health, Melbourne
Well presented. No issues Response: We sincerely thank the reviewer for the positive feedback.
Reviewer #2 Dr. Xiao Li, Case Western Reserve University
With 127 confirmed COVID-19 cases out of 1163 participants recruited in this study, the authors of this manuscript investigated the application of using wearable devices (the Ave bracelet) in detecting COVID-related physiological changes. They further developed an RNN-based machine learning approach for pre-symptomatic COVID-19 detection and showed that their algorithm could identify 68% of COVID-19 positive participants two days before symptom onset. The strengths of this study include the large sample size and the fact that all the COVID-19 cases have been laboratory confirmed. In addition, a follow-up study is currently underway, which will potentially validate this algorithm and further strength this observation. I have several suggestions regarding the revision of this manuscript which are described below.
Comment 1: Please provide more details regarding the data and data process. As the author mentioned, the bracelet saves data every 10s and requires at least four hours of relatively uninterrupted sleep. In which resolution was each parameter been collected and analyzed? Are there missing data in the data collection? If yes, what was the distribution, and how did the analysis handle the missing data? Response: 1. Resolution of collection and analysis We appreciate the reviewer for these important suggestions. We have now provided more information about the data and the data collection process throughout the manuscript. Further, the measured resolution of each physiological parameter has been added to the methods section on page 5 as follows: “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 (Figure S1).”

2. Missing data and distribution
We thank the reviewer for pointing out the lack of information about missing data and its distribution. To address these issues, we adjusted figure 4 to add missing data and distribution. Participants who did not wear their Ava bracelet for at least 29 consecutive days preceding their COVID-19 diagnosis (n=30) and participants who never paired their bracelet to the Ava app were excluded from the analysis as no bracelet data was available (n=8). Additionally, one participant (n=1) was excluded as hormone therapy may impact biophysical parameters (Figure 4).

3. Handling of missing data
We used median imputation to account for 3 cases of missing symptom duration, while no imputation was applied to the physiological parameters. As stated in our answer above, figure 4 was adjusted accordingly.

Comment 2: The authors stated that the RNN-based approach detected 68% of COVID-19 cases up to two days before SO in 66 participants with an accurate false-positive rate and laboratory-confirmed cases of SARS-CoV-2. It would be nice to show whether there are any features associated with the prediction success? For example, do the asymptomatic patients (3 patients) behave differently from those with severe symptoms or were hospitalized (11 patients)? It is probably a good idea to include more details in the model specification and algorithm performance section in the results section.

Response: We are grateful to the reviewer for this important input. We added all available known features in the results section of the manuscript, explaining the reported symptoms and the retrospectively detected asymptomatic cases. Additionally, more detailed algorithm information is now included in the section 'Model Specification and Algorithm Performance' of the results section (page 13). Due to the small sample size, we could not show any additional features in our analyses related to the prediction success.

Comment 3: According to Table 4, it seems that the sensitivity of detecting COVID was evaluated by the recall of class 1 (i.e., day 10 – day 2 before SO). What was the rationale for determining this period? What if it includes the period between day 2 to SO? Is it going to further improve the sensitivity?

Response: We thank the reviewer for highlighting this comment. Given our goals for early detection of infection to prevent transmission, we chose the -2 days cut-off based on previous work showing individuals with COVID-19 become contagious 2 days before symptom onset. We have added the rationale for this choice in the methods section on page 7 as follows: "We chose a cut-off of -2 days based on previous reports of infected participants becoming contagious two days before symptom onset [23]."

We also performed an analysis by shortening the pre-symptomatic period to -1 day prior to symptom onset. However, we could not observe an improvement in recall. This information has now also been added in the results on page 13 as follows: "Training the algorithm to detect COVID-19 one day before SO did not improve recall (data not shown)."

Minor points:
Comment 4: Probably a good idea to have consistent terms in the paper. For example, 'respiration rate' vs. 'breathing rate' were both used in the article.

Response: We agree with the reviewer and made sure to keep all terms consistent throughout the manuscript. We have, however, not found the term 'breathing rate' in our text. We have changed

‘breathing rate’ to ‘respiratory rate’ in the legend of Figure 5, including some additional adjustments to the legend.
Comment 5: Line #26, how often is the follow-up for the study? Response: We appreciate the reviewer’s suggestion to further explain and integrate the follow-up study into our manuscript. The COVI-GAPP Study, based in the Principality of Liechtenstein, is a pilot for the COVID-RED study conducted in the Netherlands (n=20,000). The COVID-RED study is testing the efficiency of the RNN algorithm developed here in real-time. As stated on page 15 in our manuscript, initial results from this larger prospective randomised study are expected in December 2022. Additional information about the COVID-RED study is now added to the abstract, methods, and discussion to make it clearer throughout the manuscript.
Comment 6: Line 45, ‘SE’ used before the first-time introduction (line 46) Response: Thank you for the correction. We have now introduced ‘symptom end’ before using its abbreviation ‘SE’.
Reviewer #3 Dr. Asma Channa, University POLITEHNICA of Bucharest
The paper proposed a methodology to detect COVID-19 patients from the data of wearable sensors (bracelet). The authors used some existing well-known machine learning algorithms to detect COVID-19 patients. Adding very small recommendations for the authors:
Comment 1: References for standardized algorithms to be added in the manuscript – SVM, k-NN etc. For instance, previously other ML algorithms have been used to detect COVID symptoms that must be included in related work. Response: We thank the reviewer for this advice. We have now referenced four additional publications that use ML algorithms to detect COVID-19: Hassantaba et al. : https://arxiv.org/abs/2007.10497 Query et al. : https://pubmed.ncbi.nlm.nih.gov/34713179/ Nestor et al.: https://www.medrxiv.org/content/10.1101/2021.05.11.21257052v1 Shapiro et al.: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7815963/
Comment 2: A brief write up must be included that discuss the reasons behind the accuracy being less than 80% Response: We thank the reviewer for bringing this missing information to our attention. We now have provided a brief explanation in the discussion section on page 15, stating the following: “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 (n=20,000). Furthermore, our investigation was based on data from individuals younger than 51 years who typically show less severe symptoms. The algorithm could perform better in older people with more severe clinical manifestations. This question will also be addressed within the framework of the COVID-RED study [13].”
Comment 3: Future work and limitations should be extended.

Response: We agree with the reviewer's recommendation. We have addressed this as follows:

1. Future work
We extended the discussion section of the manuscript with more information about generalisability, with the following information added on page 14: "This machine learning algorithm can be applied to any sensor device that measures the same physiological parameters." Further, we discuss the possibility of detecting other illnesses based on health data measured using sensor technology combined with deep learning, as follows: "In addition, detecting other illnesses using a wearable-informed machine-learning algorithm is promising [28,30]."

2. Limitations
Additional limitations are now added to the discussion section on page 15 and read as follows: "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 (n=20,000). Furthermore, our investigation was based on data from individuals younger than 51 years who typically show less severe symptoms. The algorithm could perform better in older people with more severe clinical manifestations. This question will also be addressed within the framwork of the COVID-RED study [13]. 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 (i.e. approximately 15%, depending on disease prevalence)[37,38] 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 to the current testing routine."

Reviewer #4
Dr. Nirmal Mazumder, Manipal Academy of Higher Education

The manuscript is well written and very important study at present situation. The manuscript can be accepted after minor revision.

Response: We are thankful to the reviewer for the positive feedback. We have taken the comments by the editors and reviewers into consideration and have amended our manuscript accordingly.

Comment 1: The device is well recognied and used. Did the author use the device as it is or any hardware improvement made?

Response: Our participants used the device as it is, and no hardware improvements were made. We have now amended the first paragraph on page 6 to read: "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."

Comment 2: The study is a kind of prediction model. How reliable it is ? Is it easy to use for common people?

Response: 1. Is it easy to use for common people?
As the Ava bracelet was originally developed as a medical device supporting women getting pregnant by showing their fertility window, self-use by the general population is feasible. Operating instructions and the accompanying Ava App can be found online and are accessible to the public.

2. How reliable it is?

The Ava bracelet can be used to detect COVID-19 pre-symptomatically. In our interim study using the Ava bracelet and an RNN algorithm, we were able to identify 68% of COVID-19 positive participants two days prior to symptom onset. The results section now includes additional information about the test and training phase to elaborate on the reliability. In the next step, the COVID-RED study will further improve the reliability of the Ava bracelet and the algorithm as outlined and added in the methods section on page 5.

Comment 3: What is the sensitivity and selectivity in covid detection ?

Response: Thanks to the reviewer for this question. We now address this in the abstract and results of the manuscript. The sensitivity of COVID-19 detection in the algorithm is reflected in the recall of class 1 (see table 4) and specificity in the recall of class 0 with a 70:30 train-test-split. This is now mentioned in the results section as follows: “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-tst 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 one day before SO did not improve recall (data not shown).”

Comment 4: Currently RT PCR based detection is suggested. can it be replaced by the proposed analysis?

Response: We greatly appreciate this question. If the reviewer is referring to using the Ava bracelet instead of conducting an RT-PCR test, then our answer is ‘No’. The Ava bracelet cannot replace an RT-PCR Test. The Ava bracelet may be a useful device to detect individuals developing COVID-19 in the pre-symptomatic phase. Diagnosis of SARS-CoV-2 infection, however, has to be done by RT-PCR test, which has much better diagnostic characteristics allowing to postulate the respective diagnosis. To emphasise this in our manuscript, we now state in the discussion section on page 15: “Our findings suggest that a wearable-informed machine learning algorithm may serve as a promising tool for pre- or asymptomatic detection of COVID-19. However, RT-PCR testing remains the most effective method to confirm COVID-19 infections.”

Reviewer #5
Dr. Deepa Vasireddy, Pediatric Group of Acadian

Kindly address the following:
Comment 1: Why was antibody testing chosen over RT-PCR

Response: We thank the reviewer for posing this question. The SARS-CoV-2 antibody test was not chosen over the RT-PCR test but was performed in addition to the RT-PCR test. In fact, all study participants (irrespective of COVID-19 infections) were invited twice for SARS-CoV-2 antibody tests in our study centre as stated in the methds section:
“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 assays test for pan-immunoglobulins directed against the N antigen and the receptor-binding domain of the SARS-CoV-2 spike protein [19]”.
Study participants that had COVID-19 related symptoms were additionally encouraged to be tested by RT-PCR at the Liechtenstein national testing facility. Since the RT-PCR testing was optional, this resulted in participants with positive COVID-19 confirmed by both RT-PCR and SARS-CoV-2 antibody tests and in participants with registered COVID-19 symptoms and positive SARS-CoV-2 antibody tests.

To correct this lack of clarity in our manuscript, we now state in the method section on page 6: “If participants had any symptoms during the study period, they were encouraged to visit the Liechtenstein National Testing Facility for RT-PCR testing.”

Comment 2: Can the authors comment on generalisability of these findings if other similar devices were used.

Response: We are thankful to the reviewer for highlighting this. The algorithm designed for this study will apply to other wearable devices that measure the same physiological parameters as the Ava bracelet. We added this valuable information into the discussion section on page 14, stating: “This machine-learning algorithm can be applied to any sensor device that measures the same physiological parameters.”

Comment 3: After identifying pre symptomatic individuals through such devices, are the authors trying to convey that the patients could quarantine or get tested early on as the ultimate possible utility of the study.

Response: Indeed, we fully agree with the reviewer that this is the ultimate utility of the study, and we, therefore, highlight this in the first sentence of the discussion by saying: “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 SARS-CoV-2.”
We further stated with regards to the COVID-RED study aim: “…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 [13].”

Comment 4: What are the economic costs to the patient to own such a device. False positive pick up rate and possible unnecessary testing burden with such cases?

Response: 1. Economic costs
An AVA fertility bracelet costs $279 (Ava: A Fertility Tracking Bracelet Unlike Anything Else - AvaWomen). We have mentioned this in the presentation of the bracelet on page 5 as follows: “The wrist-worn tracker is commercially available at US$ 279 and consists of three sensors that measure five physiological parameters simultaneously …”

2. False positive pick up rate and testing burden
About half of all participants were picked up with the bracelet as potentially having pre-symptomatic COVID-19 but were SARS-CoV2 negative in subsequent testing. This illustrates the point that the bracelet cannot be regarded as a tool to diagnose SARS-CoV2 infection. However, it helps to select patients for further laboratory testing, allowing for diagnosis. Half of the patients potentially having pre-symptomatic COVID-19 is considerably better than usual positivity rates (i.e. 5 to 30%) which are observed in symptomatic SARS-CoV2 patients during the pandemic. We have mentioned this in the limitations section as follows: “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 (i.e. approximately 15%, depending on disease prevalence)[37,38] 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 to the current testing routine.“

COI statements:

Reviewer: 1 Competing interests of Reviewer: None. Reviewer: 2 Competing interests of Reviewer: NA. Reviewer: 3 Competing interests of Reviewer: No competing interests. Reviewer: 4 Competing interests of Reviewer: No competing interests. Reviewer: 5 Competing interests of Reviewer: None.

VERSION 2 – REVIEW

REVIEWER	Li, Xiao Case Western Reserve University
REVIEW RETURNED	03-Apr-2022

GENERAL COMMENTS	The authors have addressed my questions/concerns.

REVIEWER	Channa, Asma University POLITEHNICA of Bucharest, Computer Science
REVIEW RETURNED	31-Mar-2022

GENERAL COMMENTS	The manuscript is clear and well written. The overall novelty of this application is sound and interesting.

REVIEWER	Mazumder, Nirmal Manipal Academy of Higher Education
REVIEW RETURNED	21-Mar-2022

GENERAL COMMENTS	The manuscript can be accepted in present form.

REVIEWER	Vasireddy, Deepa Pediatric Group of Acadiana, Pediatrics
REVIEW RETURNED	22-Mar-2022

GENERAL COMMENTS	Well conducted study. Thank you for addressing the reviewer comments and making the manuscript more comprehensible in certain areas and provision of clarifications where needed.