101005177 — COVID-RED

COVID-RED

WP7.7

D7.7 First Report from the Advisory Board

Lead contributor	1. UMCU
Other contributors	2. Ava 3. Julius Clinical 6. Sanquin

Document History

Version	Date	Description
V1.1	24.09.2021	Final version

Publishable summary

As part of the COVID-RED project, a scientific and ethical Advisory Board is set up, which will act as an independent consulting body for the management of the consortium, and provide non-binding strategic advice to the Managing Board on the scientific progress and problems that may arise during the implementation of the IMI COVID-RED initiative.

In this report, Advisory Board members have shared their questions and comments on several aspects of the study, including GDPR compliance, algorithm design, study design, participant retention, data analysis. The AB feedback was considered carefully by the consortium and the consortium responses are included in this report. The consortium will take the received feedback into account in the further course of the COVID-RED project.

Table of Contents

Introduction

COVID-RED is a public-private partnership, funded by the Innovative Medicines Initiative running from 1 July 2020 until 31 December 2021, with a possible extension into 2022. The overall goal of COVID-RED is to evaluate the use and performance of a CE-marked device (wearable), which uses sensors to measure breathing rate, pulse rate, skin temperature, and heart rate variability for the purpose of early detection and monitoring of COVID-19 in general and high-risk Dutch populations. At the same time, a mobile application will be used to track participant-reported symptoms.

A prospective, observational study is following amost 18,000 individuals from the Dutch population wearing the device and responding to participant self-report parameters via a purpose-designed app. Based on this data, an algorithm will indicate which individuals require COVID-19 diagnostic testing. To evaluate algorithm performance, the cohort will be tested for COVID-19 antibodies at the end of the study, with stored baseline samples of participants who have tested positive also being tested to determine whether the participant was already positive at baseline or was exposed to SARS-CoV-2 during follow-up.

As part of this project, a scientific and ethical Advisory Board (AB) is set up, which will act as an independent consulting body for the management of the consortium, and provide non-binding strategic advice to the Managing Board (MB) on the scientific progress and problems that may arise during the implementation of the IMI COVID-RED initiative.
This report is a first of a series of three public reports, and its aim is to share the assessment and recommendations on the scientific quality of the work conducted during the first period of the project (M1-M11).

COVID-RED Advisory Board

We set up a diverse Advisory Board (AB) specialised in the relevant scientific and ethical fields, and have thus invited six suitable experts to join the AB. The following expertise is represented in the Advisory Board: ethicist, patient advocate, representative from the Dutch National Institute for Public Health and the Environment, healthcare provider, wearable device specialist, clinical trials expert. All AB members are individuals with extensive experience, scientific and/or industrial prominence and leadership in the field of the project.

Our six appointed Advisory Board members are:

First name	Surname	Organization
Cees	Smit	Patient Advocate
Susan	van den Hof	The National Institute for Public Health and the Environment (RIVM) – Representative
Martine	de Vries	The Leiden University Medical Center (LUMC) - Ethicist
Manuel	Castro Cabezas	Julius Clinical - Healthcare provider
Björn	Eskofier	Friedrich-Alexander University – Wearable device specialist
Frank W.	Rockhold	Duke University – Clinical trials expert

Following a kick-off meeting on 2 July 2021, our Advisory Board members were asked to deliver the first review of the materials developed from month 1 to month 11 of the project, and to provide, from their expert perspective, feedback on the progress of the COVID-RED scientific initiative so far.

They were especially requested to look critical at key steps, activities, and/or missing elements in the work packages that may represent a risk for the realization of the scientific programme, and/or may affect its quality or implementation.
Ultimately, they were tasked with qualifying, according to their own expertise, the main risks attached to the scientific progress of COVID-RED, and, wherever possible, to suggest remedial actions.

The materials made available for this first review were:

- Recording and slides of the kick-off meeting, including project presentation;
- First bi-annual report (D7.4) describing the project's progress from month 1 to month 11 (July 2020 – April 2021);
- Publication of COVID-RED study protocol as available in July 2021;
- The project's website: www.covid-red.eu, as well as access to the project's Twitter account (@CovidRed) and Instagram account (covidredproject);
- Highlight of press releases about COVID-RED from month 1 to month 13:
    - Promising results from our pilot study in Liechtenstein: COVI-GAPP;
    - IMI recruitment message about COVID-RED;
    - Interview with Rick Grobbee about the project (in Dutch only);
    - Press article written by a participant, interviewing Rick Grobbee and other participants about their experiences (in Dutch only).

Advisory Board comments

General observations

The research protocol has been approved by the the ethics committee of the University Medical Center Utrecht (UMCU) on 27 January 2021.

The informed consent procedure is not explicitly described, but the website for (potential) participants of the COVID-RED study is well designed. The information and instruction videos for participants are also very informative.

The delay with the start in recruiting participants has been compensated by a high inclusion with almost 18,000 of the targeted 20,000 participants. Extensive efforts during an extended period of inclusion allowed reaching almost 90% of the target number of participants.

The study design has been well thought through, preventing and addressing potential biases as much as possible. The research protocol is very well designed and also the website for (potential) participants of the COVID-RED study is well designed. The information and instruction videos for participants are also very informative.

In general, the project is off to an excellent start. The premise to use remote monitoring to

detect disease is sound and the experimental model employed is both clever and scientifically sound. In addition, the execution to date has been impressive in terms of enrolment and randomization in the face of a constantly changing pandemic environment. Of particular note are the use of randomization, a crossover model and stratification. While none of these is novel on their own, the combination employed to study a technology rollout is unique and can serve as a model for others. To employ a rigorous study design to study a technology to better to do future trials is elegant. As for the future I think the AB is well placed to review and advise future projects.

GDPR/regulatory compliance

There is no reference to the GDPR or relevant local regulations, but I am confident that this is well organized given the approval by the UMCU ethics committee?

Consortium response

The protocol says the following under section 11.1 : “The study will be conducted in full conformance with the principles of the “Declaration of Helsinki” (64th WMA General Assembly, Fortaleza, Brazil, October 2013) or with the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the individual. The study will comply with the General Data Protection Regulation EU 2016/679 (GDPR); in Dutch: Algemene Verordening Gegevensbescherming (AVG) and the Dutch Act on Medical Devices (Besluit Medisch Hulpmiddelen).”
In addition, Appendix 2 in the protocol outlines which data is collected, how the data is processed and the relevant security measures. Upon signing the eConsent, which has been reviewed & approved by the METC of the UMCU, participants agree to their data being used for the purpose of the study. Finally, the study data is only being processed in countries that fall under the GDPR, or in countries with which the EU has an adequacy decision (e.g. Switzerland; Adequacy decisions | European Commission (europa.eu)).

Concerning the data collected by the Ava bracelet, Ava can attest that it upholds all GDPR regulations and as outlined on its website here under “Data Security”: https://www.avawomen.com/how-ava-works/healthcare/research/ As part of the full Ethics Committee review, the device and its accompanying Ava COVID-RED app were reviewed and approved by the Medische Technologie & Klinische Fysica (MTFK) department of the UMCU. Ava was required to submit an Investigational Medicinal Device Dossier (IMDD) which outlined the device’s intended use, population and design development process. Additionally, a full risk-benefit portfolio was required. A review of the device by Dutch regulatory authorities prior to recruitment start further ensured they maximized data security and minimized risk to participants.

Algorithm design – vaccination status

Due to the extended recruitment phase, most participants will be expected to have been vaccinated at the start of period 2. During the learning phase and period 1, the main share of the study population is expected to get vaccinated. If the algorithm should be different for vaccinated and unvaccinated persons, this may present a challenge for the analyses.

Consortium response

We recognize and agree with the AB that when the consortium designed the experiment, we did not account for the possibility of widespread vaccination so quickly. The rapid rollout of COVID-19 vaccines in Europe has been a truly remarkable feat, albeit one with ramifications for our analyses. We now know that detecting infections in the vaccinated continues to be very relevant because breakthrough infections are common and can be transmitted. We also know that breakthrough infections in the immunocompetent are less severe than infections in the unvaccinated: the AVA bracelet may help to alert the vaccinated that they may have an infection. However, analyses should indeed be stratified by vaccination status, which may pose a challenge for statistical power.

With the project's long term sustainability in mind, it seemed imperative that we adapt our WP deliverables to align with the real-world applications of the Ava COVID-RED app and device. Work Package 1 has considered how vaccines may impact the algorithm development and has shifted algorithm version 3's focus to handle data from vaccinated individuals. In particular, we will train a machine learning model based on data from all participants, regardless of vaccination status, retrospectively on data from across the Learning Phase, Phase 1 and Phase 2. Depending on the model parameters and specifications, this may lead to two separate models (one for vaccinated, one for unvaccinated device users) or an integrated model.
In line with this, vaccination status has been a prominent consideration when thinking through the statistical analyses that will be performed. Where appropriate and necessary analyses are stratified or corrected for vaccination status and sensitivity analyses may be performed to further investigate this effect.

Algorithm design – ethical alogorithm value

There is no formal ethics guidance / research during algorithm development (at least by an ethicist). I am curious how certain ethical challenges and questions will be tackled. First of all, will the possible effects (+/-) of this AI application as seen by relevant stakeholders (participants, society, government) be identified, not only in economic terms but also in terms of ethical values lost and gained? And will there be a reflection on which underlying values are at play? For example autonomy, privacy, medicalization, health, trust.

Consortium response

In general, WP4 will attempt to address some of the questions surrounding the value of the algorithms developed by the COVID-RED project. The AB brings up a compelling case, that perhaps an additional follow-up analysis or collaboration could query how participants and/or the broader public feel about an AI application to predict SARS-CoV-2 infection. We are currently partnering with a group of psychologists to look at ways to increase participant retention; it may be beneficial for us to explore their interest in this follow-up survey as well.

Algorithm design – bias

We know that algorithms can magnify differences and give rise to bias (see for example: https://science.sciencemag.org/content/366/6464/447.long). How will this be prevented?

Consortium response

With regard to the quesiton about bias prevention, we do not deny that algorithm development inherently leads to bias based on the decisions implemented throughout the process. We have attempted to minimize the bias, when possible. For example, although algo development was largely overseen by Ava, Work Package 1 (WP1) was involved in the review and discussion of key algo-related decision. WP1 included non-industry, device agnostic members as well as academics who work closely with other wearable device companies who rank among Ava's potential competitors. By including multiple perspectives in these conversations, WP1 sought to minimize bias related to the "black box" problem plaguing most algorithm development; given the project's design and overall structure, Ava had to be accountable to and transparent with other WP1 members, the Consortium, and its funder. Most COVID-19 publications on wearable detection of infection to date have not had as broad review by as varied stakeholders; many were either device manufacturers publishing on what their users reported (e.g., Miller et al.) or academic scholars in smaller collaborations (e.g., Shapiro et al.).

Our project will, however, struggle to minimize bias in all other areas. For example, given the Dutch population and our recruitment strategy, we will not be able to generalize our findings across racial groups. Our pilot study, COVI-GAPP, enrolled participants from Liechtenstein. This resulted in us training v1 of our algo (deployed to COVID-RED participants during the Learning Phase) on a predominantly white sample. While planned revisions to the algorithm will incorporate data from Dutch participants, nevertheless, we will not be able to generalise our model's performance to other ethnic groups. While some wearables struggle to detect biological signals on darker skin pigments, we know from internal analyses at Ava that the device measures changes in physiological parameters equally well across races. Thus, biases due to racial differences in this trial will stem from the recruited population's ethnic make-up; short of having a longer recruitment period with race-based quota sampling, we could not have corrected for this up front. Even with a balanced racial make-up in our sample, residual bias would have likely arisen from the disparities inherent in who catches SARS-CoV-2. Studies in the US have demonstrated that black and Latin individuals are more likely to contract the virus, given their comparative overrepresentation in service jobs that require they interact in person with others. We have noted these potential reasons for discrepancies in test positive rate and will be conducting secondary analyses to better understand whether we see similar disparities in our sample.

The algorithms employed in COVID-RED show additional bias based on where we conducted the clinical trial. They rely on testing requirements outlined by the Dutch health authorities. In particular, an automatic red alert is provided to participants if they report in-app that they have a symptom which would qualify them for test according to the national health ministry's guidelines. Applying these same algorithms to samples drawn from other countries would highlight the biases inherent in the algorithm's design.

We have tried to control for other sample-based biases when possible, recruiting from a broad range of ages and having a relatively equal break down in genders, to be able to demonstrate our algorithm's generalizability to those populations. We agree that sub-analyses of algorithm performance are merited, to help provide insight into whether additional areas of bias may nevertheless exist. Only one other paper on COVID-19 detection using wearable devices to date has examined their algorithm's performance by race and gender (see Nestor et al.). We appreciated their thorough review of potential biases and plan to do the same. In addition, several of the secondary/exploratory objectives will also quantatively dive deeper into the

(independent) predictiveness of certain features of the algorithm and thus will give a greater understanding of what drives the algorithm performance.

Algorithm design – "pingdemic" situation

Is there any chance that this algorithm, when used broadly, will lead to a "Pingdemic" situation as in the UK? How will this be tackled?

Consortium response

With regard to the Pingdemic situation, the problem in the UK was caused by having to go into quarantine after receiving an exposure notification. That is not the case with AVA bracelet notifications, especially not in the vaccinated. The Ava COVID-RED app provides indications about potential deviations in physiological parameters. It is not designed or approved from a regulatory perspective for diagnostic purposes nor to provide medical advice. Instead, it alerts the user to when physiological parameters deviate from their own personal baseline and gives suggestions for possible behavioral changes. When the participant reports a symptom that could be related to COVID-19 but is not included on the Dutch health ministry's list of testable symptoms, the user receives a yellow indication. This alert suggests the user consider self-isolating and checking in with the app the next day. In the case of a red alert, the user is told that their "physiological parameters and/or symptoms that might be related to a potential COVID-19 infection. Please consider seeking testing and/or input from a medical professional." The wording was carefully chosen so users could not misconstrue it as medical advice or diagnostic test result. It is left to the user's discretion how they respond to the health indication. The COVID-RED protocol asks participants to get tested if they receive a red alert but does not require they quarantine. We do not forsee a similar "Pingdemic" occurring as a result of the COVID-RED clinical trial.

Study design – virus variants

Will the different variants from the virus which are developing so rapidly, all be detected by the antibody assay in the lab? Can new variants of COVID-viruses quickly be implemented in testing of the participants during the study? Has this been covered sufficiently? How is it possible to differentiate between antibodies generated by vaccination from those by infection? The published protocol mentions that that is possible, but since no methodology has been added, an external reviewer cannot test its validity.

I wonder whether the study team has considered a follow-up study for those infected during the study period, plus a sample of uninfected participants, to allow for continued measurements among those infected with and without long COVID symptoms. These usually are defined as still having symptoms at least 3 months after infection, so would require an extension.

Consortium response

The SARS-CoV-2 virus produces several proteins. The spike (S1) protein, which sticks out of the viral membrane, is the most immunogenic. All COVID-19 vaccines that are used in the Netherlands only stimulate the production of antibodies against the viral S1 protein;

antibodies against other viral proteins are absent. In a natural infection, we would expect the patient to make antibodies against multiple viral proteins, including for example the nucleocapsid (N) protein. This is why we are using N protein detection to differentiate between natural infection and vaccination. New variants of the virus do alter their spike proteins, but so far, the changes have been small, which means that the vaccines and the antibody detection assays still work.
In practice, we had to slightly adjust our outcome measurement plan due to the mass vaccine role out happening quicker than expected. We managed to rapidly respond to this by using anti-N serology tests for all serology samples taken after the baseline sample for those who indicated being vaccinated for COVID-19.

Participant retention

The retention of participants will be the biggest challenge in this project. With over 10% discontinuations having occurred so far, efforts to retain most current participants are pivotal. Although the planned initiatives were discussed during the first meeting, this issue deserves continuous attention. As the number of withdrawals or lost to follow-up is sizable, it needs to be investigated further. It may be interest to see if this is related to randomized group and/or sequence.

Retention could become a problem because of the start of the vaccination programmes – will people still be motivated after they have received their vaccination? – but on the other hand, the emergence of new variants of COVID-viruses could very well compensate for this.

Consortium response

We fully recognize how crucial retention of the recruited participants is for the study as well as the challenging determinants that may affect this. For the statistical analysis of this study it is crucial that we have as many participants as possible remaining in the study. On the other hand, dropout or non-compliance is inevitable with this being a fully decentralized study and social, political, and societal factors all heavily influencing the behaviour of the participants during follow-up. These unique conditions have also made it extremely difficult to predict the amount of dropout or non-compliance to be expected in this study.

Obviously, all possible actions have been taken to encourage compliance and keep participants in the study. While designing the registration/onboarding process, efforts have been made to make this process as smooth as possible, while still requiring sufficient effort in order to actually be randomized into the study. Other actions include regularly sending newsletters on pertinent topics such as to keep participating after vaccination, responding to any questions or concerns coming into the helpdesk, regularly sending participants feedback on their compliance, along with associated medal (gold, silver, bronze), etc. In addition, we recently also investigated which demographic characteristics and other factors such as vaccination status contributed to dropout that we experienced so far to be able to identify subgroups that we might target for further reminders. In summary, despite the challenging circumstances we are trying as best we can, within the limits of the study budget, to keep participants informed of their progress and the progress of the study and remind them that it is essential that they keep contributing. This all serves to keep the participants engaged and compliant during the course of the study, which has been a challenge in itself given the constantly changing landscape in which this study is being conducted.

With regard to the relevance of the project, digital interventions that support source- and contact-tracing (the AVA bracelet could be considered a source-tracing intervention) are more important in this phase of the epidemic than they were during lock-downs. Society is opening up in the absence of sterilising herd immunity, which means that R will increase (even with the current Delta variant; even more so if a more infectious variant is introduced). While most people are vaccinated, the proportion of unvaccinated or poorly immunised is still sufficiently high to overwhelm the healthcare system once again, unless we stay on top of the source- and contact-tracing. We might want to communicate this to the participants. Many people think that vaccination alone will get us there and that is not true.

Data analysis – incidental findings

Is it possible that the project results in incidental findings and if so, how are these incidental findings then handled?

Consortium response

Without knowing more about the types of incidental findings referred to by the AB, we can speak only broadly about the planned analytic process. The consortium has identified a set of primary analyses related to the main research question and the algorithm's performance, which we will tackle first. Additional secondary analyses (as outlined in the published protocol and statitistical analysis plan) will look at sub-groups in the data, as well as other questions around time to first alert and health care utilization. We recognize some interesting, previously unforseen findings may arise; Work Package 3 is tasked with the overall data analysis and has been having ongoing discussions regarding emergening exploratory analyses (e.g., algo performance based on vaccine status). As outlined in Work Package 5's Description of Action, the consortium is also dedicated to ensuring the data's long-term accessibility according to FAIR data principles. We will make the data available to third-party researchers after the project's end, who may have interesting additional questions they want to investigate which the consortium never considered. We hope that this will allow for the full exploration of incidental findings and anticipate many COVID-RED collaborators will continue to analyze sub-sections of the data for years to come.

Data analysis – social and psychological determinants

It is interesting to read that some participants see value in the vital sign measurements outside of the aim of detecting infections in an early stage. Good that these messages are used to optimize retention.
From the press clippings, it seems that social and psychological determinants play a role in compliance that participants didn't anticipate before (for instance, alcohol use records in the app). Is the intention to make an overview of these social and psychological determinants in using AI-technology with an aim to improve further use of these technologies or to identify them as possible barriers for its further use?

Consortium response

Model-based and analytic considerations drove the inclusion of self-report features in the Ava

COVID-RED app's daily diary. We know from prior research that some substances, like alcohol, can impact the measurement of physiological features. Because the COVID-RED algorithms were designed to detect deviations from baseline physiological parameters, we needed to be able to parcel out the variance in parameter measurements due to participant behavior. By asking participants to report on potentially confounding variables in their daily diary, we could implement a better specified model that provides fewer false positive alerts. The algorithm will still provide health indications to the user, even in the absence of data about the user's behavior. While we have requested COVID-RED participants complete this daily, they are not penalized for withholding the information. The only exception to this is the reporting of COVID-19 symptoms, upon which the symptom-only algorithm condition relies.

The intention of including other social and psychological determinants (e.g., alcohol intake) was to improve the performance of the device's AI-technology; their impact on participant retention has become an interesting, albeit unanticipated, consequence necessitating further consideration when thinking through the device's long term sustainability and real-world use.

Data analysis - study power analysis

I have nowhere seen a power analysis, nor how the expected drop out will influence the power of the study?
The power to detect differences in sensitivity between the two measurement methods (with and without vital signs) will be highly dependent on the number of endpoints (SARS-CoV-2 infections during period 2 and retention. Both are difficult to predict.

Consortium response

Section 4.4 of the protocol as well as the Statistical Analysis Plan give an indication of the expected power of the primary analyses under certain scenarios as expected at the beginning of the study. The power calculation was highly dependent on the amount of infections to be expected during follow-up which was extremely difficult to predict given the influence of governmental responses to the pandemic as well the rapid mass rollout of the vaccination campaign. In addition, given the difficulty in predicting the percentage of dropout to be expected in this study, as described in the responses to the AB questions on participant retention above, if was very difficult to take this into account for the power calculations.

We agree with the AB members and despite our best efforts, non-compliance and dropout will have a large influence on this trial. Moreover, we will have to stratify by vaccination status, which may negatively influence our power. This will be thoroughly investigated and considered when executing the statistical analysis.