PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence, pathophysiology, prediction and health-related quality of life of long COVID: study protocol of the longitudinal multiple cohort CORona Follow Up (CORFU) study
AUTHORS	Ghossein-Doha, Chahinda; Wintjens, Marieke SJN; Janssen, Emma BNJ; Klein, Dorthe; Heemskerk, Stella; Asselbergs, Folkert; Birnie, Erwin; Bonsel, Gouke J.; van Bussel, Bas; Cals, Jochen; Ten Cate, Hugo; Haagsma, Juanita; Hemmen, Bena; van der Horst, Iwan; Kietselaer, Bastiaan; Klok, Frederikus; de Kruif, Martijn; Linschoten, Marijke; van Santen, Susanne; Vernooy, Kevin; Willems, Loes H.; Westerborg, Rosa; Warle, Michiel; van Kuijk, Sander

VERSION 1 – REVIEW

REVIEWER	Francesca Bai University of Milan, Department of Health Sciences
REVIEW RETURNED	25-Jul-2022

GENERAL COMMENTS	The authors present the study protocol of a multicenter cohort study investigating the prevalence of long COVID at 2 years. The outcomes are well described, methods are clear and detailed and seem appropriate. Furthermore, the research question is still interesting; the better characterization of PASC also in the long term period and the identification of a prediction model could be extremely useful in clinical practice. Thus, I think that the study protocol is suitable for the publication in the current form.

REVIEWER	Onur Boyman University Hospital Zurich, Department of Immunology
REVIEW RETURNED	03-Sep-2022

GENERAL COMMENTS	In this manuscript titled "Prevalence, pathophysiology, prediction and health-related quality of life of long COVID: design of the longitudinal multiple cohort CORona Follow Up (CORFU) study", Ghossein-Doha, Wintjens and colleagues report the design of a longitudinal and multiple cohort COVID-19 follow-up study, accronymized CORFU. The aim of CORFU is to study the prevalence, pathophysiology, prediction and health-related quality of life of long COVID. To this end, the authors intend to aggregate seven COVID-19 cohorts of the Netherlands and assessed long COVID patients at 3, 6, 12, 18, and 24 months after infection, compared to individuals that did not suffer from COVID-19. The primary outcome of the study will be the prevalence of long COVID at 2 years after infection. Secondary outcomes will include health-related quality of life, physical functioning, and the prevalence of thromboembolic complications, respiratory complications,

	cardiovascular disease, and endothelial dysfunctioning. Long COVID is a serious and prevalent condition. The CORFU, as proposed by the authors, would be a very useful study to address many of the central questions related to long COVID. The manuscript is interesting to read and could be improved by addressing the following points. Major comments 1) How will the authors ensure that the group of individuals that did not suffer from COVID-19 indeed consists of subjects that did not have contact with SARS-CoV-2 prior to their study? Several articles have shown that PCR-based detection of SARS-CoV-2 in nasal swabs, saliva or other specimens is limited to only a few days after infection and, similarly, antibody-based assays of blood (or serum) samples are insufficient to reliably detect all individuals that had contact with SARS-CoV-2 and developed asymptomatic or very mild symptoms. 2) The authors suggest that "CORFU findings may be used to inform national and international guidelines on diagnostics, treatment and follow-up of long COVID and contribute to developing a (new) more accurate long COVID definition, likely differentiating long COVID phenotypes.". The authors should elaborate on what they foresee as accurate long COVID definition and which long COVID phenotypes they hypothesize to be studying. Minor comment a) The definition used to identify long COVID was not provided or not easy to find in the manuscript. The authors should provide this information.

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1
C2. The authors present the study protocol of a multicenter cohort study investigating the prevalence of long COVID at 2 years. The outcomes are well described, methods are clear and detailed and seem appropriate. Furthermore, the research question is still interesting; the better characterization of PASC also in the long term period and the identification of a prediction model could be extremely useful in clinical practice.
Thus, I think that the study protocol is suitable for the publication in the current form.

A2. We kindly thank reviewer 1 for the time and effort to critically read and review our manuscript.

REVIEWER 2
Q3. How will the authors ensure that the group of individuals that did not suffer from COVID-19 indeed consists of subjects that did not have contact with SARS-CoV-2 prior to their study? Several articles have shown that PCR-based detection of SARS-CoV-2 in nasal swabs, saliva or other specimens is limited to only a few days after infection and, similarly, antibody-based assays of blood (or serum) samples are insufficient to reliably detect all individuals that had contact with SARS-CoV-2 and developed asymptomatic or very mild symptoms.

A3. We thank the reviewer for this comment and understand the concern related to the group of individuals that did not suffer from COVID-19. In the CORFU study, we only use self-reported questionnaires. Therefore, some differential misclassification may occur in cases that had no or only very mild symptoms, never got tested, and will likely report never or not sure having suffered from

COVID-19. The control group might therefore be slightly biased. In asking participants whether they suffered from COVID-19 we distinguish between “yes; confirmed by test”, “possible; symptoms, but not confirmed by test”, and “no”. This enables to reduce bias by excluding the possible COVID-19 cases. The remaining misclassified controls are expected to having suffered from very mild disease.

Due to the presumed association between COVID-19 severity and persisting symptoms1,2, we hypothesize that extremely mild cases are not those that are likely to develop long COVID. However, if they do, they may slightly decrease the difference between long COVID cases and controls and hence, result in slightly conservative estimates of the total burden of disease. We do not have the

possibility to ad-hoc test participants of this large control cohort, so effects of this will be discussed in any manuscript that (also) uses data of these controls. We would like to kindly point out to the reviewer that the majority of our study aims will be studied using former COVID-19 cases, diagnosed using the diagnostic criteria employed at the time of diagnosis.

However, during the second and third wave of the pandemic, nationwide public health policy included routinely testing of the entire population through PCR-, rapid antigen- and self-tests, whether or not having symptoms. We expect that this test policy had diminished the proportion of missed asymptomatic cases during this specific time period in the COVID-19 pandemic. In order to give more context to the CORFU study, we added Supplementary Table S2 to the Appendix of the manuscript. This table gives a detailed description of the Dutch national lockdown, testing policy and vaccination strategy timeline from February 2020 onwards. In the manuscript we referred to this Appendix by adding the following (subsection Data collection: CORFU questionnaire, page 8): “As study participants were included at different time point in the COVID-19 pandemic, depending on their date of first infection, different contextual factors might apply such as lockdowns, the availability of testing material and testing policy, and the vaccination strategy at that time. These factors are presented in detail in Supplementary table S2.”

We have elaborated more on the limitation regarding our control group in the discussion section of the manuscript (final paragraph, page 13), by adding the following: “This might result in some misclassification of cases that had no or only very mild symptoms, never got tested, and will likely report never or possibly having suffered from COVID-19. We hypothesize that these extremely mild cases are not those that are likely to develop long COVID, but if they do, they may slightly decrease the difference between long COVID cases and controls and hence, result in slight conservative estimates of the total burden of disease.”

Q4. The authors suggest that "CORFU findings may be used to inform national and international guidelines on diagnostics, treatment and follow-up of long COVID and contribute to developing a (new) more accurate long COVID definition, likely differentiating long COVID phenotypes.". The authors should elaborate on what they foresee as accurate long COVID definition and which long COVID phenotypes they hypothesize to be studying.

A4. We thank the reviewer for this insightful comment. Currently, the definition of the World Health Organization (WHO) is used in most studies to report long COVID. However, this definition is very broad and does not specify which and how many symptoms should be considered to diagnose long COVID. In addition, symptoms are currently only attributed to long COVID when they are not ascribed to an alternative diagnosis. Consequently, similar symptoms in individuals may not be attributed to long COVID due to comorbidities in which overlap in symptoms occurs. As a result it is challenging to operationalize long COVID in population-based and clinical studies. Being more specific and potentially add clinical parameters (e.g. biomarker, imaging, etc.) might enhance the application of the definition in research, but especially clinical practice. Besides, long COVID symptoms are very heterogeneous, probably covering a variety of underlying pathophysiological processes. We expect

that distinguishing subtypes (i.e. phenotypes) of long COVID-19 might be required in order to develop and provide tailored, effective therapy strategies.

At present, we are in the data-collection phase of our study. Therefore, we are not able yet to predict what phenotypes of long COVID will be detected, though we might expect differentiation based on (multi-)organ systems being involved for example the respiratory, cardiovascular or neurological system. However, our approach will be independent of a-priori hypotheses. For the unsupervised clustering algorithms (i.e. algorithms for which no outcome is used as is in regression modeling) that will be used, we will select symptoms and complaints from all domains (e.g. respiratory, mental health, etc.) after a data-reduction step. The results of this analysis, which has not been performed before, may well have an impact on the current working definition of long COVID.

We have clarified this by briefly adding the following to discussion section of the manuscript (third paragraph, page 12): “Besides, the current WHO long COVID definition remains broad and unspecific, thereby lacking accurate differentiation of its heterogeneous appearance into clinical phenotypes. Defining such phenotypes with potentially adding clinical parameters (biomarkers, imaging, etc.) might enhance clinical workability, and thereby diagnostics, and the development of tailored therapies based on underlying pathophysiology.”

C5. The definition used to identify long COVID was not provided or not easy to find in the manuscript. The authors should provide this information.

A5. We thank the reviewer for this comment. Initially, the WHO definition of long COVID will be used to report long COVID. As our study design is suitable for studying longitudinal data, possible fluctuations in symptoms over time might be unraveled. Besides, after potential identification of long COVID phenotypes these will also be further considered.

The WHO definition of long COVID is already described in the introduction section of our manuscript (first paragraph, first sentence, page 5). In the manuscript we further clarified that the WHO definition will be our working definition, by adding the following specification to the methods section (subsection Outcome variables on page 8): “Initially, the WHO definition will be used to define long COVID. Potentially identified long COVID phenotypes as part of WP1 and other international developments within the field will also be further considered throughout the study.”

VERSION 2 – REVIEW

REVIEWER	Onur Boyman University Hospital Zurich, Department of Immunology
REVIEW RETURNED	08-Oct-2022

GENERAL COMMENTS	This is a revised version of the manuscript titled "Prevalence, pathophysiology, prediction and health-related quality of life of long COVID: design of the longitudinal multiple cohort CORona Follow Up (CORFU) study", Ghossein-Doha, Wintjens and colleagues. My previous major comments were: 1) How will the authors ensure that the group of individuals that did not suffer from COVID-19 indeed consists of subjects that did not have contact with SARS-CoV-2 prior to their study? Several articles have shown that PCR-based detection of SARS-CoV-2 in nasal swabs, saliva or other specimens is limited to only a few days after infection and, similarly, antibody-based assays of blood (or serum) samples are insufficient to reliably detect all individuals that had

	contact with SARS-CoV-2 and developed asymptomatic or very mild symptoms. The authors responded: "Thirdly, especially in the first COVID-19 wave (March 1 – June 30, 2020) for non-hospitalized patients, not all suspected COVID-19 cases were tested due to capacity and test-material constraints in the Netherlands. However, these patients were included in (some of) the cohorts despite the lack of a confirmed infection. Therefore, findings might be based on suspected instead of confirmed infections. The same holds for controls not being tested due to the absence of symptoms (Table S1). This might result in some of cases that had no or only very mild never got and will likely never or possibly having suffered from COVID-19. We hypothesize that these mild cases are not those that are likely to develop long COVID, but if they do, they may slightly decrease the difference between long COVID cases and controls and hence, result in slight conservative of estimates of the total burden of disease." Based on reports in the literature, I have to disagree with the authors' response and assumption "that these mild cases are not those that are likely to develop long COVID". This is and remains a major shortcoming of the planned study, which the authors should address in a more robust and satisfactory manner. For example, the authors could limit their analysis to those individuals that were tested positive for SARS-CoV-2, either by PCR or serology. 2) The authors suggest that "CORFU findings may be used to inform national and international guidelines on diagnostics, treatment and follow-up of long COVID and contribute to developing a (new) more accurate long COVID definition, likely differentiating long COVID phenotypes.". The authors should elaborate on what they foresee as accurate long COVID definition and which long COVID phenotypes they hypothesize to be studying. The authors responded: "Besides, the current WHO long COVID definition remains broad and unspecific, thereby lacking accurate differentiation of its heterogeneous appearance into clinical phenotypes. Defining such phenotypes with potentially adding clinical parameters (biomarkers, imaging, etc.) might enhance clinical workability, and thereby diagnostics, and the development of tailored therapies based on underlying pathophysiology." Unfortunately, the authors' response is unspecific and imprecise, and it does not address my previous point. The authors should elaborate on which long COVID phenotypes they expect to observe and what their hypotheses are, thus this reviewer and the future readers can judge whether their study protocol is adequate to reach their goals. The authors have adequately addressed my previous minor point.

VERSION 2 – AUTHOR RESPONSE

REVIEWER 2

Q1. Based on reports in the literature, I have to disagree with the authors' response and assumption "that these mild cases are not those that are likely to develop long COVID". This is and remains a major shortcoming of the planned study, which the authors should address in a more robust and satisfactory manner. For example, the authors could limit their analysis to those individuals that were tested positive for SARS-CoV-2, either by PCR or serology.

A1. We thank the reviewer for responding to our revision of the CORFU study protocol manuscript. When responding to the reviewer on October 6, 2022, we might have partly misinterpreted the reviewer's initial comment. We hope our answer below clarifies the methods used in the CORFU study.

The aims of the CORFU study will be studied based on data from CORFU participants who suffered from COVID-19. The majority of these participants are included in cohorts with COVID-19 patients who were admitted to the hospital ward and/or Intensive Care Unit or who presented to the Emergency Department. Hence, the majority of the included CORFU participants suffered from COVID-19 which was confirmed, during hospital admission, by either a positive PCR test for SARS-CoV-2 and/or a positive CT scan of the chest based on the COVID-19 Reporting and Data System (CO-RADS) Score (score 4-5 by a radiologist). Only a subgroup of CORFU participants (likely) suffered from COVID-19 at home based on self-reported questionnaires (i.e. unspecified positive test or the presence of COVID-19-related symptoms). However, as this is self-reported, we are not able to confirm whether these participants have truly been infected with SARS-CoV-2 in the past. To be fully transparent to future readers, we specified the method section of the manuscript (subsection Participants, page 6): **"The study population consists of Dutch (former) COVID-19 survivors and non-COVID-19 controls, who have been included in one of the cohorts. Former COVID-19 cases are either confirmed by a positive PCR test for SARS-CoV-2 and/or a positive CT scan of the chest based on the COVID-19 Reporting and Data System (CO-RADS) Score (score 4-5 by a radiologist), or likely based on self-reported questionnaires (i.e. unspecified positive COVID-19 test or the presence of COVID-19-related symptoms). The study population is categorized into five subgroups:**

- **Patients who suffered from confirmed COVID-19 admitted to the hospital ward;**
- **Patients who suffered from confirmed COVID-19 admitted to the ICU;**
- **Patients who suffered from either confirmed or likely (i.e. self-reported) COVID-19 at home;**
- **Patients who suffered from confirmed COVID-19 and needed inpatient or outpatient rehabilitation after infection at home or in the hospital (ward and/or ICU);**
- **Controls who (likely, i.e. self-reported) did not suffer from COVID-19.".**

The primary analyses of the CORFU study aims will be performed with data from CORFU participants who have suffered from either confirmed or likely COVID-19. Sensitivity analyses will be performed for the subgroup of participants who have suffered from confirmed COVID-19 (i.e. positive PCR test for SARS-CoV-2 and/or a CO-RADS Score of 4-5). This was specified in the method section of the manuscript (subsection Work packages and data analysis, page 9): **"The primary analyses of the WP1-3 aims will be performed with data from CORFU participants who (likely) suffered from COVID-19. Sensitivity analyses will be performed for the subgroup of participants who suffered from confirmed COVID-19 (i.e. positive PCR test for SARS-CoV-2 and/or a CO-RADS Score of 4-5).".**

And in the discussion section (page 13-14): **"Thirdly, especially in the first COVID-19 wave (March 1 – June 30, 2020) for non-hospitalized patients, not all suspected COVID-19 cases were tested due to capacity and test-material constraints in the Netherlands. For CORFU, this means that there is a lack of confirmed infections for participants of the community-based POPCORN cohort who self-reported to have (likely) suffered from COVID-19 based on an unspecified positive test or the presence of COVID-19-related symptoms. In order to limit the impact of this limitation, the primary analyses will be repeated in the sensitivity analyses on the subgroup of COVID-19 cases with a confirmed infection.".**

With regards to the group of individuals that did not suffer from COVID-19, the reviewer is correct that we are not completely sure whether this group consists of subjects that did not have contact with

SARS-CoV-2. This limitation is, however, unavoidable as we have access to the data from a source population of 3.293 Dutch community-based participants who have already completed three rounds of questionnaires during the course of 2020 to 2022 during the COVID-19 pandemic. Unfortunately, no serology data were collected in this group of participants.

It is important to note that, as mentioned before, CORFU study aims will be studied based on participants who suffered from COVID-19, diagnosed using the diagnostic criteria employed at the time of diagnosis. Data from individuals that most likely did not suffer from COVID-19 will be used as reference in order to compare the presence and severity of long COVID symptoms and health-related quality of life in COVID-19 cases with the general Dutch population to help quantify burden of disease of long COVID. As already described in the discussion section of the paper (page 13), this is a crucial comparison currently lacking in the majority (79%) of long COVID research, but required to identify and quantify attributable symptoms objectively. Despite the limitation, we feel that taking into account these data is an important addition to the available literature, and we will transparently report the limitation in our current study protocol manuscript and in future publications on CORFU study findings. In the current manuscript we altered the paragraph in our discussion section (page 14): "**The same holds for controls from the POPCORN cohort who (likely) did not suffer from COVID-19: there is a possibility that these controls did suffer from COVID-19, but that they for example did not test due to the absence of symptoms or test-material constraints or that their tests were false negative not being tested due to the absence of symptoms (Supplementary Table S1). This might result in some misclassification of cases that had no or only very mild symptoms, never got tested, and will likely report never or possibly having suffered from COVID-19. This will be described when reporting CORFU study results, and, when deemed relevant, additional (stratified) analyses will be conducted.".**

Lastly, we understand why the reviewer disagrees with our assumption that mild cases are not those that are likely to develop long COVID. Available evidence in the literature is inconclusive and, hence, we removed this sentence from the manuscript (discussion section, page 13) and think the altered paragraph (discussion section, page 14, as described above) is a clear depiction of this CORFU study limitation.

Q2. Unfortunately, the authors' response is unspecific and imprecise, and it does not address my previous point. The authors should elaborate on which long COVID phenotypes they expect to observe and what their hypotheses are, thus this reviewer and the future readers can judge whether their study protocol is adequate to reach their goals.

A2. The current long COVID definitions do include a variety of symptoms, but do not discriminate between subgroups of patients with similar patterns in occurrence and intensity of symptoms. We hypothesize that these different subgroups exist and, if proven so, may have a substantial impact on working definitions of long COVID. One of the goals of the CORFU study is to estimate these 'phenotypes' (i.e. subtypes: patients with similar expression of long COVID symptoms). Subdividing the heterogeneous long COVID syndrome into phenotypes may guide towards patient-tailored therapy strategies.

We acknowledge that a priori hypotheses may exist on what phenotypes might be found. However, we think that hypothesis-driven analyses could hamper multi-dimensional distinction of long COVID phenotypes, which might provide a more accurate reflection of reality. Important phenotypes may depend on combinations of previously reported domains of symptoms (e.g. respiratory, cardiovascular), but subtypes may also depend on previously unidentified combinations. Therefore, in defining phenotypes of long COVID we will follow a data-supporting approach independent of a-priori hypotheses, namely unsupervised clustering algorithms as described in our previous response. This

has not been performed yet for long COVID in literature. Long COVID phenotypes will be based on clusters of patients based on their scores on the long COVID symptom items of the questionnaire. Clusters will be estimated using unsupervised machine learning techniques, which are not confirmatory (of prior hypotheses) in nature. Our study provides ample precision to estimate the number and characteristics of these clusters based on our large sample, but will also make it possible to perform sensitivity analyses by stratifying cluster analyses to subgroups of former COVID-19 patients (e.g. patients in the Intensive Care unit). This approach could result in clusters involving multiple organ systems based on frequency associations.

In order to make this more clear in the manuscript, we altered the method section (subsection Work packages and data analysis, page 10): **"Furthermore, long COVID 'phenotypes' (i.e. subtypes: patients with similar expressions of symptoms) will be estimated. Important phenotypes may depend on combinations of previously reported domains of symptoms (e.g. respiratory, cardiovascular), but phenotypes may also depend on previously unidentified combinations. Clusters of patients will be estimated using unsupervised machine learning techniques with K-means and hierarchical clustering, which are data-supportive and thereby not confirmatory (of prior hypotheses) in nature.".**

VERSION 3 – REVIEW

REVIEWER	Onur Boyman University Hospital Zurich, Department of Immunology
REVIEW RETURNED	04-Nov-2022

GENERAL COMMENTS	The authors addressed my points.