Statistical Analysis Plan for PROVID-LD

WHO COVID-19 - Evaluation of the Efficacy of Probiotics to Reduce the Occurrence of Long COVID (PROVID-LD)

NCT number: NCT05080244

Analysis Plan Author: Mi-Suk Kang Dufour, Epidemiologist/Statistician

Analysis plan reviewers: Salomé Bouskila, Statistician

Version 1.0

Last Update: October 31th, 2023

1. Table of Contents

Section 1: Administrative Information

2. Title and trial registration

WHO COVID-19 - Evaluation of the Efficacy of Probiotics to Reduce the Occurrence of Long COVID (PROVID-LD)

Official Title: Evaluation of the efficacy of probiotics taken during the acute phase of COVID-19 to reduce the occurrence of long COVID

Registered with Clinicaltrails.gov

NCT number: NCT05080244

Trial registration is maintained by Sarah Bilodeau

3. Protocol Version

This document was written based on information contained in study protocol version 9 dated 2021-10-04.

4. Description of this document

This document describes the statistical analysis plan (SAP) for the PROVID-LD trial. The SAP was developed according to guidelines found in ICH E91. This SAP contains more technical and detailed elaboration of principal features of the analysis describes in the protocol, and includes procedures for executing the statistical analysis of the primary and secondary variables. Planned ancillary studies will not be included in this SAP and will have their own separate SAP. Others unplanned exploratory analyses not identified in the SAP will be clearly identified in the reporting as well as any deviation in the analyses from this original SAP.

The SAP should be executed once the trial database lock (DBL) has been achieved. Analyses will be double-blinded for the randomization group. All the analyses will follow the Statistical Analysis Plan (SAP) described in this document. Analyses will be done under the responsibility of the trial statistician at the Data Coordinating Center (Applied Clinical Research Unit (URCA), CHU-Sainte-Justine, Montreal).

During the analysis processes, the study Principal Investigator, the Executive Committee and the Steering Committee will be kept blind to the randomization group until final disclosure of the study results.

5. Statistical analysis plan revisions

Any changes to the SAP will be logged in the following table:

Protocol version	Updated SAP version	Section number	Description of and reason for change	Date changed

6. Roles and responsibilities

Chercheurs

Université de Sherbrooke

Jean-Charles Pasquier, MD, PhD, principal investigator
Jean-François Beaulieu, PhD, immunologie et biologie cellulaire, co-investigator
Felix Camirand Lemyre, PhD, statistiques, co-researcher
Mélissa Généreux, MD, MSc, santé publique, co-researcher
Isabelle Laforest-Lapointe, PhD, biologie, co-researcher
Mélanie Plourde, MD, PhD, gériatrie, co-researcher
Antoine Lewin, PhD, statistiques, co-researcher
Alain Piché, MD, MSc, infectiologie, co-researcher
Louis-Charles Fortier, PhD, microbiologie, collaborator
Paul Farand, MD, MSc, cardiologie, collaborator
André Cantin, MD, MSc, pneumologie, collaborator
Alfredo Menendez, PhD, immunologie et infections bactériennes, collaborator
Nadine Sauvé, MD, médecine, collaborator
Gérald Gahide, MD, radiologue, collaborator

Université Laval

Nils Chaillet, PhD, méthodologie, co-investigator
Guy Boivin, MD, PhD, infectiologie, collaborator
Julie Robitaille, PhD, nutrition, collaborator

CHU Ste-Justine

Benoît Mâsse, PhD, Biostatistics, Lead Biostatiscian, co-researcher
Mi-Suk Kang Dufour, Statistician
Salomé Bouskila, Statistician

Partenaires collaborateurs

Lallemand Solutions Santé

Marie-Laure Oula, MSc
Sylvie Binda, PhD

Coordination Centre de recherche du CHUS

Sarah Bilodeau, MSc.
Olivier Fortin, Data manager

Authorized and approved by Trial steering and management committee

Name	Signature	Date

Section 2: Introduction

7. Background and rationale

Probiotics may be considered as an option of treatment for long COVID since they have anti-viral effect, trigger immunomodulation and have low side-effects. This randomized controlled trial aims to reduce the number of patients with long COVID by 25% 90 days after the COVID-19 diagnosis by taking probiotics in a symptomatic population, self-caring at home. During the acute phase of the disease, participants will take two capsules (probiotics or placebo) per day for 10 days and one capsule (probiotics or placebo) per day for the following 15 days. A follow-up will be done twice during the acute phase, 14 days and 28 days after starting to take the investigational product (compliance to treatment, side effects, etc.). At inclusion and at Day14, Day30 and Day90 after the COVID-19 diagnosis, a questionnaire will be administered (COVID-19 symptoms, anxiety, functioning difficulties, etc.) and 2 saliva (microbiota analyzes) self-samples will be performed.

COVID-19 disease caused by a new coronavirus (SARS-CoV-2) has received worldwide attention. No effective treatment against COVID-19 is available and the current vaccination appears essential to control the pandemic. However, many questions remain, such as the durability of immunity and the appearance of variants. While the acute phase has been widely studied, the long-term consequences are unknown, and a new burden is emerging: the long COVID. The long COVID is defined by the persistence of symptoms of the disease for more than 4 weeks after diagnosis. Several strong arguments support the study of probiotics for COVID-19: 1) probiotics act on viruses by various well described mechanisms (reduction of absorption, cellular internalization of the virus, production of metabolites/substances having a direct antiviral effect and immunomodulation); 2) probiotics are considered with a high level of evidence (meta-analysis) to reduce diarrhea (associated with antibiotics and Clostridium difficile) and for respiratory tract infections and 3) probiotics are affordable and available with low side-effects.

8. Study Objectives

Primary objective: Reduce by 25% the number of patients with LONG-COV during follow-up at D90 by taking probiotics during the acute phase of COVID-19.

Secondary objectives (Obj S):

Objective S1: Compare the proportion of patients presenting COVID-19 symptoms and severity of symptoms at D14, D30 and D90 by taking probiotics during the acute phase of COVID-19.

Objective S2: Describe the symptoms severity and evolution by study group and baseline characteristics.

Objective S3: Determine/identify prognostic factors measured at baseline (inclusion) associated with LONG-COV (sociodemographic, clinical factors).

Section 3: Trial Methods

9. Trial design

Inclusion Criteria:

- 18 years and over
- ≤ 10 days between the COVID-19 diagnosis and the inclusion
- Having symptoms of the COVID-19 at inclusion
- Living in Quebec for the next 90 days
- Self-caring at home
- Able to take medication alone
- With access to a phone or to the Internet
- Able to give informed consent

Exclusion Criteria:

- Taking probiotic supplements at inclusion
- Taking antibiotics for a reason other than COVID-19 at inclusion
- Allergies to soy, lactose, yeast, maltodextrin, vitamin C, potato starch, magnesium stearate, hypromellose or titanium dioxide
- Has a chronically weakened immune system (AIDS, lymphoma, chemo-radio-corticosteroid therapy, immunosuppressive pathology)
- Has any immunosuppressive medication (Anti-rejection treatment following an organ transplant)
- Was treated with chemo-radio-corticosteroid therapy in the last 6 months
- Has active cancer
- Already taking part in another randomized clinical trial
- Pregnant or planning to become pregnant; breastfeeding or will be breastfeeding in the next few months

10. Randomization

Randomization with a 1:1 ratio, probiotics vs. placebo, in the first line immediately after inclusion. Prior to the start of the study, the randomization list will be generated by the statistician considering age and gender stratification. Age stratification will be done in 4 strata: 18-29 years; 30-49 years; 50-69 years and 70 years and older.

Randomization will be done via the REDCap application. The statistician will create different randomization tables according to gender (female/male) and age strata. These tables will then be imported into REDCap for automatic randomization based on the age and sex information of each participant.

During the course of the trial and until the database is locked, study personnel at the site (principal investigator and research team) and at the Sponsor (research specialists and laboratory manager) will not be informed of the group allocation. Based on a blinded randomization method (according to the algorithm built into the REDCap platform), participants will be assigned to one of the 2 groups. The CHUS pharmacy will prepare the products according to the allocation on REDCap. Lallemand Health Solutions will deliver all products before the beginning of the study to the CRCHUS pharmacy.

11. Level of statistical significance

Because LONG-COV is a new disease, few data are available to determine the proportion of patients with persistent symptoms or the occurrence of symptoms associated with LONG-COV at D30 or D90. To be conservative in the size calculation, we used an expected proportion of 50% in the placebo group at D90. Because the procedure is over-the-counter and inexpensive, we calculated the size to detect a moderate relative difference of 25% (50% versus 37.5%) with a two-sided alpha of 5% and power of 80%. 494 patients (247 per group) are required. Considering a 20% attrition at D90, a total of 618 patients will have to be recruited.

Final analyses will begin in September of 2023 after database lock.

Unless otherwise specified, the alpha will be set at 0.05 for statistical significance.

12. Adverse events reporting

All effects experienced by participants for the duration of the study (from randomization to 30 days after the last dose) must be documented (noting duration, intensity, severity, predictability, causality). All unexpected events (not listed above) that are potentially related to intervention and considered serious must be reported within 24 hours: to the ethics committee, to Health Canada and to Lallemand Health Solutions. A full report should be sent within 7-15 days, depending on the seriousness of the situation.

13. Definition of adherence and how adherence will be presented

Adherence will be estimated via self-reported adherence in the logbooks at D14 and D28 completed online or on paper (according to the preference of the participant). If at least 80% of the doses prescribed was taken i.e. 20 pills out of 25, we consider that the participant adhered to the protocol.

14. Lost to follow-ups

We initially anticipated that 20% of participants will be lost to follow-up, according to the 1st PROVID-19 pilot study statistics. Indeed, 21.4% were lost to follow-up prior to 90 days.

15. Definition of analysis populations e.g. Intention-to-treat, per-protocol

Intent to treat analyses will include all participant randomized to either control or intervention arm including participants lost to follow up.

Per-protocol analyses will include all participant randomized to either control or intervention arm, excluding participants lost to follow up or those with less than 80% adherence to the assigned intervention.

16. CONSORT diagram

Per recommendations in the CONSORT guidelines, the statistician will generate a flow chart indicating the number of individuals screened and enrolled, the reasons for ineligibility, the number of individuals randomized, and the number of individuals in each arm with data at each time point. Some participants have incomplete forms mid-study but were not lost to follow-up, hence the differences in Totals.

Approached
n = (*Approx. 2000*)

Not screened	n = (23)
Contact information missing	7
Couldn't be reached	7
Participant removed because didn't meet criteria of selection	1
Incomplete form	1
Refusal to participate	3
No data collected	12

Screened
n = (738)

Not eligible	n = (55)
Immunosuppressive medication	0
Allergies to soy, lactose, yeast, maltodextrin, vitamin C, potato starch, magnesium stearate, hypromellose or titanium dioxide	5
Has received chemo/radio/corticotherapy within the last 6 months	1
Has an active cancer	1
Has a chronically weakened immune system	4
Is pregnant; plans to become pregnant; is breast-feeding or will be breast-feeding within the next few months	8
Is under 18 years of age	3
Does not have at least one of the Covid-19 symptoms listed in the Quebec Self-Care Guide	17
Patient has another condition that would prevent safe participation in the study	4
The first positive(+) screening test was not performed within the last 10 days	8
Does not have access to a telephone or the Internet	7
Not able to provide informed consent	7
Not domiciled in Quebec for the next 90 days	8
Is not self-medicating	7
Not in home care	7
Already participating in another randomized clinical trial	0
Currently taking antibiotics for reasons other than Covid-19	5
Currently taking probiotic supplements	10

Eligible
n = (683)

Not participants	n = (65)
Anticipation of lack of time to complete all follow-up steps	9
Anxiety about taking oral medication (pills)	1
No interest	4
Fear of side effects	1
Need time to think	1
Too complicated/demanding	9
Too long	3
Hypoglycemia problem	1
Uncomfortable taking probiotics	1

Randomized
n = (618)

	Completed (at least partially)	Lost to follow-up
At Baseline	618	0
At D14	523	91
At D30	502	23
At D90	479	25
Total	**479**	**139**
Total %	77.5	22.5

Intervention
n = (307)

	Completed (at least partially)	Lost to follow-up
At Baseline	307	0
At D14	261	43
At D30	251	12
At D90	240	12
Total	**240**	**67**
Total %	78.2	21.8

Placebo
n = (311)

	Completed (at least partially)	Lost to follow-up
At Baseline	311	0
At D14	262	48
At D30	251	11
At D90	239	13
Total	**239**	**72**
Total %	76.8	23.2

Statistical Analysis

17. Descriptive analyses and positivity checks

Distribution of sociodemographic and clinical variables measured at inclusion will be described for each group. The proportion of missing data will be described for each variable.

The probability of intervention conditional on baseline variables will be calculated for all observations to confirm the effectiveness of randomization and to verify the balance of baseline factors in each intervention arm. Overall probability will be graphed to confirm that there are no violations of positivity. Should there be positivity violations, the statistician will investigate and characterize participants who have a less than 5% or greater than 95% probability of intervention. The statistician will then notify the investigators and discuss whether a restriction of the analysis is necessary. Potential covariates associated with the outcome will be evaluated and considered in the analysis to limit the confusion bias.

To assess the quality of participant blinding, participants assessment of their arm will be compared against the true randomization status.

A table describing the demographic and clinical characteristics at baseline randomization will be generated, however statistical testing will not be done comparing individual variables between groups.

Participation in follow-up assessments and completeness of outcome data will be calculated and compared for each time point by study arm. A sensitivity analysis on loss to follow-ups will be performed using weights or extreme case scenarios. In the latter method, we will assume that all losses to follow-up had or did not have LONG-COV at D90, enabling us to estimate the potential bias.

18. Intention to treat

For the primary analyses, we will follow an intention to treat model. All participants will be assigned intervention status based on their randomized assignment. We will use an inverse probability of censoring for missing outcome data. Participants lost to follow-up will be censored.

19. Adherence per protocol analysis

Per-protocol analyses will include all participant randomized to either control or intervention arm, excluding participants lost to follow up or those non-adherent to the assigned intervention. Adherence will be estimated via self-reported adherence in the logbooks at D14 and D28 completed online or on paper (according to the preference of the participant). If at least 80% of the doses prescribed was taken i.e. 20 pills out of 25, we consider that the participant adhered to the protocol.

20. Long-term COVID-19 (LONG-COV)- Outcome definition

Background
Based on the WHO, long-term COVID is defined as follows: “It is defined as the continuation or development of new symptoms 3 months after the initial SARS-CoV-2 infection, with these symptoms lasting for at least 2 months with no other explanation. »
(https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition)

Definition of COVID-LD in PROVID-LD
Long-term COVID-19 (LONG-COV) status was defined according to three criteria:

A Binome of COVID-19 symptoms from FLU-PRO:
A binome is a recurrent symptom, i.e. present at D30 and D90. We selected participants with at least one symptom recorded at D30 and D90 with a value of moderate level or more ([2:4]).
Also, we selected the participants who presented the symptoms of "Loss of Taste" and/or "Loss of Smell" at D30 or D90.
A Binome of symptoms from the Quality of life indicators:
We reviewed the participants who showed a moderate level or more ([2:3]) at D30 and D90 for difficulty performing current activities, pain/comfort and anxiety/depression.
Worsening of Functioning difficulties:
For the functioning difficulties, participants were asked to estimate their functional status before COVID-19 in the form at D30 and in the form at D90. The smaller value was retained if an inconsistency was detected between these two values. We harmonized the recorded values pre-COVID-19 at D30 and D90. We selected participants with a non-zero value at D30 and a difficulty value at D90 higher than pre-COVID-19 for Concentration and memory, and the ability to Take care of themselves.

The status of FLU-PRO and quality of life indicators before COVID-19 disease was initially unknown. Thus, we have operationalized the WHO’s definition to select patients. If at least one of the criteria was met, the participant would be qualified for an expert’s review. Every diagnosis based on Binomes was reviewed by expert on LONG-COV physicians. Those initially selected who potentially qualified for LONG-COV were called after the study to verify if symptoms were present or not before the positive detection of COVID-19. The final decision was on the expert physician responsibility.

First, initial analyses were carried out using a binary LONG-COV criterion (yes/no). Secondly, LONG-COV cases were categorized according to their level of severity and symptoms profile over time. Participants in the intervention and the placebo arms of the trial were compared.

21. Descriptive analyses of outcome measures

Distribution of score measures will be calculated for each of: COVID-19 symptoms (FLU-PRO and loss of smell and/or taste), quality of life indicators (GAD-7) and functioning difficulties (WG-SS) at D30 and D90 timepoints. The distributions will be generated overall and comparatively by intervention arm using t-test or rank sum testing as appropriate based on the distributions.

Individual symptom frequency for those items included and not included in the FLU-PRO instrument score will be described and compared between groups using chi-squared or fishers exact tests as appropriate.

22. Stratifying variables

The analysis will follow the randomization strategy. Gender and age were used as stratifying variables. We will use the same categories in the analysis.

23. Primary outcome analyses

The proportions of participants with LONG-COV at D90 will be compared using the Cochran-Mantel-Haenszel test, adjusting for stratification.

24. Secondary outcome analysis

S1: The presence and severity of symptoms at D14, D30 and D90 will be explored by doing a descriptive analysis for each study-time. The proportions of participants with LONG-COV at D14, D30 and D90 will be compared using the Cochran-Mantel-Haenszel test, adjusting for stratification. A table will summarize the proportion of each symptoms at each time with the means severity per study arm.

S2: A GEE analysis (with linear or logit link function depending on the variables) will use Symptoms Severity as the outcome over all time points and include baseline characteristics. This is to assess whether the intervention reduces the number and severity of symptoms. A faster decrease is expected with intervention.

We will supplement these analyses with a GEE (with a logit link function), using a binary LONG-COV status as the outcome and including measurements at Baseline, D14, D30 and D90 and indicator variables for stratification.

If there are differences in censoring by group, an additional GEE model will be used with inverse probability of censoring weights (IPCW) to assess the bias potentially induced by attrition during follow-up to D90. This method first estimates the probability of dropout to construct IPCWs, and then fits the GEE model with these IPCWs.

S3: Finally, logistic regression will be used to determine potential prognostic factors at baseline associated with LONG-COV at D90. The relationship between other variables including self-reported health, and LONG-COV will be assessed.

25. Adverse and Serious adverse events

Adverse events (AE), including serious adverse events (SAE), will be monitored and tracked using a logbook. They will be coded according to MedRA standards and will be summarized per intervention group.