medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

TITLE

Prevalence and risk factors for long COVID after mild disease: a longitudinal study with a symptomatic control group.

AUTHORS

- Ana Beatriz C Cazé– ana.caze@ufba.br - Universidade Federal da Bahia, Salvador, Bahia, Brazil
- Thiago Cerqueira-Silva, MD - csilvathiago@hotmail.com - LIB and LEITV Laboratories, Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil; Universidade Federal da Bahia, Salvador, Bahia, Brazil
- Adriele Pinheiro Bomfim, M. Sc., - adrielepbomfim@gmail.com – Instituto Gonçalo Muniz, Fiocruz, Salvador, Brazil; Universidade Federal da Bahia, Salvador, Bahia, Brazil
- Gisley Lima de Souza - gisleysouza19.1@bahiana.edu.br - Escola de Medicina e Saúde Pública da Bahia; Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil
- Amanda Canário Andrade Azevedo, MD, M.Sc. – amandacanario@gmail.com - Hospital Santa Izabel, Santa Casa da Bahia, Salvador, Bahia, Brazil
- Michelle Queiroz Aguiar Brasil, MD, M. Sc. - michelleqas@gmail.com - Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil; Federal University of Bahia, Salvador, Bahia, Brazil
- Nara Rúbia Santos - narabusiness@hotmail.com - Departamento de Vigilância em Saúde/Vigilância Epidemiológica, Campo Formoso, Bahia Brazil
- Ricardo Khouri, PhD - ricardo_khouri@hotmail.com - Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Clinical and Epidemiological Virology, Belgium; Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil

- Jennifer Dan, MD, PhD – jdan@ucsd.edu - Center for Infectious Disease and Vaccine Research, la Jolla Institute for Immunology (LJI); Department of Medicine, Division of Infectious Disease and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA 92037, USA
- Antonio Carlos Bandeira, MD, PhD, – antoniobandeira@gmail.com - UNI-FTC-Faculdade Tecnologia Ciências Medical School, Bahia; Central State Laboratory LACEN-Bahia; Infectious Diseases Advisor
- Luciano Pamplona de Goes Cavalcanti, MD, PhD - pamplona.luciano@gmail.com - Universidade Federal do Ceará, Fortaleza, Ceará, Brazil; Centro Universitario Christus, Fortaleza, Ceará, Brazil
- Manoel Barral-Netto, MD, PhD - manoel.barral@fiocruz.br - LIB and LEITV Laboratories, Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil; Universidade Federal da Bahia, Salvador, Bahia, Brazil
- Aldina Maria Prado Barral - aldina.barral@fiocruz.br - Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Bahia, Brazil; Instituto de Investigação em Imunologia, São Paulo, São Paulo, Brazil
- Cynara Gomes Barbosa*, PhD - cynara.barbosa@gmail.com - Universidade Federal da Bahia, Salvador, Brazil
- Viviane Sampaio Boaventura *, MD, PhD, - vsboaventura@gmail.com - LIB and LEITV Laboratories, Instituto Gonçalo Moniz, Fiocruz, Salvador, Bahia, Brazil; Universidade Federal da Bahia, Salvador, Bahia, Brazil

*Those authors equally supervised this study

Corresponding author

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Viviane Sampaio Boaventura

Email address: vsboaventura@gmail.com

Telephone: + 55 71 31762211

Postal address: 121 Waldemar Falcão Street, Candeal, Salvador, Bahia, Brazil, CEP 40296

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Background

There is limited data on the prevalence and risk factors for long COVID, with a shortage of prospective studies with appropriate control groups and adequate sample size. We therefore performed a prospective study to determine the prevalence and risk factors for long COVID.

Methods

We recruited patients age ≥ 15 years who were clinically suspected of having acute SARS-CoV-2 infection from September 2020 to April 2021. Nasopharyngeal swabs were collected for RT-PCR 3-5 days post symptom onset. Clinical and sociodemographic characteristics were collected using structured questionnaires from persons positive and negative for SARS-COV-2. Follow-up was performed by telephone interview to assess early outcomes and persistent symptoms. For COVID-19 cases, 5D-3L EuroQol questionnaire was used to assess the impact of symptoms on quality of life.

Results

We followed 814 participants (412 COVD-19 positive and 402 COVID-19 negative persons) of whom the majority (741 / 814) had mild symptoms. Both the COVID-19 positive and the COVID-19 negative groups had similar sociodemographic and clinical characteristics, except for the rate of hospitalization (15.8% vs 1.5%, respectively). One month after disease onset, 122 (29.6%) individuals reported residual symptoms in the COVID-19 positive group or the long COVID group versus 24 (6%) individuals in the COVID-19 negative group. In the long COVID group, fatigue, olfactory disorder, and myalgia were the most frequent symptoms which occurred in the acute phase. Compared to recovered patients, female sex, older age and having > 5 symptoms during the acute phase were risk factors for long COVID. Quality of life was evaluated in 102 out of 122 cases of long COVID with 57 (55.9%) reporting an impact in at least one dimension of the EuroQol 5D-3L questionnaire.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Conclusion

In this prospective study consisting predominantly of patients with mild disease, the persistence of symptoms after acute disease was highly associated with long COVID-19 (29.6% vs 6% of COVID negative group). The risk factors for long COVID were older age, female sex, and polysymptomatic acute disease.

KEYWORDS

COVID-19; SARS-CoV-2; long COVID; frequency; risk factor; outpatients; post-COVID syndrome; persistent symptoms; post-Acute COVID Syndrome; PASC.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

BACKGROUND

Long COVID has received less attention from the scientific community than acute COVID-19, representing less than 2% of COVID-19 publications in PubMed Clinical Queries, August 31, 2022). Several important aspects remain undefined such as its prevalence. Long COVID has been reported to effect between 7.5% to 89% of patients [1, 2, 29, 30]. Most of these studies do not include well-matched controls which is important because long COVID-symptoms are nonspecific and may be attributed to other factors such as the physical and mental impact of having COVID-19 infection or hospitalization [3,4]. Prospective studies enrolling symptomatic individuals with and without confirmed COVID-19 are necessary to better estimate the prevalence of long COVID and its associated risk factors.

Long COVID is anticipated to have a large impact on the healthcare system [5]. Identifying risk factors for developing long COVID will help mitigate its impact. Some risk factors have been suggested for long COVID; however, these studies consisted of previously hospitalized patients [1] or were retrospective analyses from electronic health records of not hospitalized patients (31). We therefore performed a longitudinal study of patients with COVID-like symptoms who tested positive or negative for SARS-CoV-2 to assess the prevalence and risk factors for long COVID and its impact on the quality of life.

METHODS

Study population

We recruited individuals age ≥ 15 years with clinical suspicion of acute SARS-CoV-2 infection in health care units from three municipalities of Bahia State-Brazil (Irecê, Campo Formoso, and Lauro de Freitas). Individuals who had difficulty in reporting symptoms, those who were > 20 days post symptom onset at the time of recruitment, or those who reported

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

SARS-CoV-2 re-infection were excluded. The study was conducted from September 2020 to April 2021. The ancestral strain or gamma variant of SARS-CoV-2 was dominant in Brazil from November 2020 to August 2021 (**Additional file 1: Fig S1-2**) [6].

Exposure measurement

All patients were tested for COVID-19 by nasopharyngeal RT-PCR. Only individuals who had an RT-PCR in the first ten days of disease onset were considered eligible for the follow-up. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) protocol was followed.

Recruitment strategy and data collected

Consecutive in-person recruitment included the application of a structured survey (Questionnaire 1) using Research Electronic Data Capture (REDCap) software hosted by the Gonçalo Moniz Institute (IGM/FIOCRUZ) in Bahia, Brazil. The following information was collected: sociodemographic characteristics (age, gender, skin color, body mass index, anthropometric data, municipality of residence), comorbidities (hypertension, diabetes, heart disease, chronic lung disease, other), date of symptom onset, acute symptoms (fever, myalgias, arthralgias, cough, shortness of breath, nasal congestion, sneezing, coryza, sore throat, sinus pain, retroocular pain, ear pain, headache, chest pain, chest discomfort, rash, abdominal pain, loss of appetite, vomiting, diarrhea (2 or more episodes in the last 24 hours), weakness, fatigue, lightheadedness or fainting, loss of smell, and loss of taste).

Follow-up and outcome measurement

SARS-CoV-2 positive (COVID-19) and negative (non-COVID-19) patients who reported COVID-like symptoms were followed. Negative cases were recruited on the same

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

day as positive cases before nasopharyngeal RT-PCR. The date of symptom onset was considered Day zero. A second structured questionnaire (Questionnaire 2) was conducted 30 to 60 days after disease onset by telephone to document the RT-PCR result, acute symptoms, and clinical outcome. Participants were classified as non-hospitalized or hospitalized in the intensive care unit (ICU) or non-ICU). To avoid memory bias, questionnaires 1 and 2 were applied in a short interval of time.

A third structured questionnaire was conducted 60 to 250 days after disease onset by telephone to evaluate residual symptoms (Questionnaire 3). Questions included residual or new symptoms: cough, fatigue, shortness of breath, headache, chest pain, dysphonia, dysphagia, loss of appetite, loss of smell, loss of taste, myalgias, arthralgias, fever, and other non-listed symptoms. Patients reporting suspected or confirmed re-infection were excluded from the final analysis. A schematic diagram of the study is provided (**Additional file 1**: **Fig S3**).

Outcomes of interest

Long COVID was defined according to the Centers for Disease Control and Prevention classification [7]. Persistence of symptoms was defined as the presence of at least one residual symptom > 1 month post disease onset. The list of residual symptoms presented on the questionnaire was based on literature [8,9].

The European Quality of Life 5 Dimensions 3 Level (EQ-5D-3L) was conducted via telephone to assess health dimensions: mobility, self-care, daily activities, pain/discomfort and anxiety/depression. The EQ-visual analogue scale (VAS) was also conducted with 0 rated as worst health imaginable and 100 as the best imaginable health [10].

Statistical analysis

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Categorical data were described using absolute values and percentages/frequencies. Continuous data were described using mean (standard deviation, SD) and median (interquartile range, IQR). The correlation matrix was constructed based on the frequency of persistent symptoms reported by the patients. Multivariable logistic regression was performed to explore risk factors related to long COVID, applying the odds ratios (OR; 95% CI). Analyses were done in R (version 4.2.0) [26].

RESULTS

In total, 1,268 patients with suspected SARS-CoV-2 infection were recruited and answered the first questionnaire 4 days (range of 3-5 days) post symptom onset. After excluding cases outside the RT-PCR recommended window period, reported re-infection, or those lost to follow-up, 814 (64.2%) patients were considered eligible for the analysis of residual symptoms (**Fig. 1, Additional file 1: Fig. S3**). Questionnaires 2 and 3 were conducted 40 (31-59) days and 102 (63-150) days post symptom onset, respectively. No difference in sociodemographic data and frequency of positive RT-PCR were observed between the included (n=814) and excluded (n=454) participants **(Additional file 1: Table S1**). Among the included participants, 412 tested positive (COVID-19 group) and 402 tested negative (non-COVID-19 group) for SARS-CoV-2. Demographic characteristics for the COVID and non-COVID groups were as follows, respectively: a median age of 36 (IQR) and 34 (IQR) years, female gender of 55% and 64%, mild disease in 84% and 99% (**Table 1**). More patients were hospitalized in the COVID-19 group compared to the non-COVID-19 group (15.9% vs 1.5%, respectively, **Table 1**). Follow-up time was 3.3 months for the COVID group and 3.7 months for the non-COVID group (**Additional file 1: Fig. S4**).

After one month post disease onset, per the CDC definition for long COVID (September 16, 2021), 122/412 (29.6%) individuals from the COVID group reported residual

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

symptoms and were classified as long COVID cases, while the remaining 290 individuals were fully recovered. From the non-COVID group, residual symptoms were reported by only 24 (6%) individuals (**Table 1**). Among long COVID cases, 57 (13.2%) presented with more than two residual symptoms.

Within the COVID group, fatigue (56, 13.7%), olfactory disorder (42, 9.9%), myalgia (36, 8.7%), gustatory disorder (27, 6.6%), and headache (26, 6.1%) were the residual symptoms most frequently reported (**Fig. 2**). These symptoms started mainly during the acute phase. Other symptoms not reported during the acute phase, such as memory loss and hair loss, were reported by the patients with long COVID (**Additional file 1: Table S2**). Compared to those who recovered, the long COVID group had more women (63.1% vs. 52.1%) and older individuals [40 (32-51) vs. 35 (28-47) years] (**Table 2**). Hospitalization for acute disease was comparable between long COVID (21/122, 17.2%) and recovered patients (44/290, 15.2%).

The risk of long COVID increased with age and was significant for persons 50 years of age or older (OR 2.44, 95% CI 1.29-4.66). Individuals with > five symptoms in the acute phase (OR 3.15, 95% CI 1.37-8.55) and females (OR 1.55, 95% CI 0.99-2.44) were more likely to develop long COVID (**Table 3**).

Co-occurrence of long COVID symptoms was also observed. Fatigue was reported by at least 50% of those that complained of gustatory dysfunction, anorexia, dysphonia, chest pain, headache, breathlessness, or myalgia. Moreover, olfactory dysfunction was reported by 80% of participants with gustatory dysfunction (**Additional file 1: Fig. S5**).

Finally, we have evaluated the impact of long COVID on the quality of life. The EQ-5D-3L questionnaire was administered to 102 out of 122 patients. Fifty-seven patients (55.8%) reported alteration in at least one dimension of QoL, and 41 patients (41.1%) had an impact in the dimension of pain and depression (**Additional file 1: Table S3**).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

We have followed RT-PCR positive and negative patients with COVID-like symptoms from the first days post symptom onset to 63-150 days post symptom onset to determine the prevalence of long COVID-19. Long COVID occurred in almost one-third of test-positive cases, the majority with mild acute disease, compared to 6% of negative patients with residual symptoms, supporting the association between SARS-CoV-2 infection and residual symptoms. The risk of long COVID increased with age > 50 years old, female gender, and the development of > five symptoms during the acute phase. Together these findings demonstrate that residual symptoms after mild disease were mainly associated with COVID-19 with well-defined demographic and clinical risk factors.

The proportion of patients with COVID-19 who developed long COVID has varied from 7.5% to 89% [1, 2, 28, 29]. Most studies only included patients who tested positive for SARS-CoV-2. One strength of our study is the inclusion of a comparator group of symptomatic patients, recruited in the same time/location, who tested negative by SARS-CoV-2 RT-PCR. Additionally, most symptoms investigated in the follow-up questionnaires were evaluated at the time of recruitment, in a blinded manner, prior to the SARS-CoV-2 test result. The inclusion of a control group is important considering that many residual symptoms attributed to long COVID, such as fatigue and headache, are nonspecific and could be triggered or aggravated by stress and psychosocial problems connected to the global health crisis [11-13]. Indeed, we observed that non-confirmed COVID-19 cases reported similar residual symptoms after acute illness, albeit at a much lower frequency. Few studies have included a control group composed of symptomatic patients that tested negative for SARS-CoV-2 [14, 30]. In a cohort of healthcare workers, olfactory disorders and hair loss were related to long COVID while positive and negative cases both had similar frequencies of

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

exhaustion/burnout and fatigue [14]. In a population-based digital retrospective cohort including outpatients and well-matched controls, 62 symptoms were associated with long COVID [30]. The inclusion of a control group is even more valuable for hospitalized cases, considering the increased risk of sequelae associated with iatrogenic procedures for treating severe disease [15]. Together, these findings highlight the importance of including a control group when determining the frequency of sequelae to avoid overestimating long COVID cases.

We observed a similar proportion of long COVID in groups with mild and moderate/severe disease [16,17]. We enrolled a small number of severe COVID-19 cases and therefore could not conclude an association between disease severity and sequelae. Severe COVID-19 increases the risk of long COVID by 1.2 to 8x among hospitalized cases compared to those not hospitalized [18]. Nonetheless, the finding that long COVID occurs in a significant proportion of individuals after mild disease highlights its potential impact on the healthcare system, as COVID-19 variants and sub-variants continue to spread worldwide.

Fatigue was the most frequent sequelae and occurred isolated or associated with other clinical manifestations. As previously reported [19], we also observed a vast impact of long COVID on the quality of life. Although the mechanism is unknown, an exacerbation of the immunological response and multisystemic involvement [20] seems to be involved. Studies investigating the pathogenesis of long COVID are essential to improve treatment and attenuate the impact.

Similar to previous studies [3, 9, 17, 21-23, 30], we found that long COVID was more prevalent in women and could be attributed to a greater propensity for women to go to a clinic [22]. Indeed, females represented 55% of the COVID group and 64% of the non-COVID group in our study. Sex hormone differences in immune response [24] and autoimmunity triggered by SARS-CoV-2 have been suggested to be involved in the disease

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

pathogenesis of long COVID [23]. Other studies have found that individuals with more symptoms during the acute phase were at a higher risk for developing long COVID [9, 14]. There are conflicting results regarding age in the risk for developing long COVID. In a cohort of healthcare workers in Switzerland, the majority of whom had mild disease, young age was associated with a higher risk for developing long COVID [14], in contrast to our findings. Additional prospective studies including multicenter cohorts with larger sample sizes and older patients should be performed to evaluate if the risk for developing long COVID differs among age groups, racial groups, and ethnicities. Prospective studies to develop predictive models and algorithms should be performed for the early detection of long COVID cases.

Long COVID has different definitions. We have used the CDC's definition for long COVID (September 16, 2021), that includes persistence of symptoms after four weeks of the onset of disease. Recently, a WHO-coordinated Delphi consensus established a period of more than three months after COVID-19 onset for the case definition of long COVID [25]. In our study, only 445 of the calls were performed three months after disease onset. By this definition, long COVID was detected in 68/216 (31.5%) of our patients. Residual symptoms were also observed in 17/229 (7.4%) of the non-COVID-19 patients. As observed in our study, the prevalence of long COVID may vary between studies in the absence of a universal consensus to define long COVID [4]. This impacts the estimation of the prevalence of long COVID cases worldwide.

This study has some limitations. The loss to follow-up was around 36% of patients and may have influenced the estimates of residual symptoms. Similar demographic and clinical characteristics were observed among patients that were included and the ones that were lost to follow-up. In addition, most participants were young, with mild disease, which does not permit generalizing these findings to the elderly and those with severe disease.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Moreover, the participants were aware of their diagnosis of COVID-19 when questionnaire 3 was applied, which may have led to an overestimation of persistent symptoms. Finally, the short time to follow-up did not allow for an analysis of the duration of long COVID.

CONCLUSIONS

Long COVID was detected in 29.6% of patients with mild COVID-19 disease, with older age, female sex, and polysymptomatic acute disease as the main risk factors for persistent symptoms. Estimating the long COVID prevalence is important for preparing the healthcare system to assist and guide these patients. Further studies should be conducted to evaluate the impact of different variants of SARS-CoV-2 on long COVID and the potential impact of vaccines in reducing residual symptoms.

LIST OF ABBREVIATIONS

STROBE – Strengthening the Reporting of Observational Studies in Epidemiology

RT-PCR – Reverse Transcriptase Polymerase Chain Reaction

REDCap – Research Electronic Data Capture software

IGM/FIOCRUZ – Gonçalo Moniz Institute

ICU – Intensive Care Unit

SD – Standard Deviation

IQR – Interquartile Range

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

DECLARATIONS

Ethics approval and consent to participate: Informed consent was obtained from all participants. The Ethics in Research Committee of the Gonçalo Moniz Research Center approved this study (approval no. 4.315.319/202).

Consent for publication: Not applicable

Availability of data and materials: The dataset generated and/or analyzed during the current study are not publicly available due to legal and ethical restrictions. Dataset are available upon request.

Funding: This study was supported by MCTI/CNPq/FNDCT/MS/SCTIE/Decit (nº 07/2020) MB-N, AB and VSB are research fellows from CNPq, the Brazilian National Research Council. T.C-S. is a PhD student at the Post-Graduation Program in Health Sciences-UFBA, which is supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil, finance code 001.

Competing interests: The authors do not have conflicts of interests to declare for this work.

Authors' contributions: A.B.C.C. and V.B structured the original draft. V.B. and C.B. were responsible for conceptualization and methodology. A.P.B., G.L.S, A.C.A.A, M.Q.A.B, N.R.S and A.B.C.C collected data from the patients. T.C.S made the formal analysis statistics. A.M.P.B. and V.B. participated in the project administration and in the funding acquisition. TCS, JD, ACB, LPGC, MBN, ABN, CGB substantively revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements: We wish to thank all patients who were willing to participate in the study. In addition, we would like to thank Hospital Aeroporto and Municipal Health Secretariats of Irecê and Campo Formoso for logistical support and to Roberta de Carvalho Freitas for patient-related support.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

[1] Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global Prevalence of Post COVID-19 Condition or Long COVID: A Meta-Analysis and Systematic Review. The Journal of Infectious Diseases. 2022 Apr 16;jiac136.

[2] Nasserie T, Hittle M, Goodman SN. Assessment of the Frequency and Variety of Persistent Symptoms Among Patients With COVID-19. JAMA Network Open. 2021 May 26;4(5):e2111417.

[3] Amin-Chowdhury Z, Ladhani SN. Causation or confounding: why controls are critical for characterizing long COVID. Nature Medicine. 2021 Jun 17;27(7):1129–30.

[4] Alkodaymi MS, Omrani OA, Fawzy NA, Shaar BA, Almamlouk R, Riaz M, et al. Prevalence of post-acute COVID-19 syndrome symptoms at different follow-up periods: a systematic review and meta-analysis. Clinical Microbiology and Infection. 2022 Feb;28(5).

[5] Kelly BD, Gulati G. Long COVID: the elephant in the room. QJM: An International Journal of Medicine. 2021 Nov 26;115(1):5–6.

[6] Nagib W. Home [Internet]. Genomahcov - Fiocruz. Available from: http://www.genomahcov.fiocruz.br/

[7] CDC. COVID-19 and Your Health [Internet]. Centers for Disease Control and Prevention. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/

[8] Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Scientific Reports [Internet]. 2021 Aug 9 [cited 2021 Aug 16];11(1):16144. Available from: https://www.nature.com/articles/s41598-021-95565-8

[9] Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. Nature Medicine [Internet]. 2021 Apr 1 [cited 2021 May 8];27(4):626–31. Available from: https://www.nature.com/articles/s41591-021-01292-y

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

[10] EuroQol. EQ-5D User Guides – EQ-5D [Internet]. EQ-5D User Guides. 2022 [cited 2022 Sep 17]. Available from: https://euroqol.org/publications/user-guides

[11] Bartoszek A, Walkowiak D, Bartoszek A, Kardas G. Mental Well-Being (Depression, Loneliness, Insomnia, Daily Life Fatigue) during COVID-19 Related Home-Confinement—A Study from Poland. International Journal of Environmental Research and Public Health. 2020 Oct 12;17(20):7417.

[12] Usher K, Bhullar N, Jackson D. Life in the pandemic: Social isolation and mental health. Journal of Clinical Nursing. 2020 Apr 6;29(15-16).

[13] Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The Psychological Impact of Quarantine and How to Reduce it: Rapid Review of the Evidence. The Lancet. 2020 Feb 26;395(10227):912–20.

[14] Strahm C, Seneghini M, Güsewell S, Egger T, Leal O, Brucher A, et al. Symptoms compatible with long-COVID in healthcare workers with and without SARS-CoV-2 infection – results of a prospective multicenter cohort. Clinical Infectious Diseases. 2022 Jan 28;75(1).

[15] Evans RA, Leavy OC, Richardson M, Elneima O, McCauley HJC, Shikotra A, et al. Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. The Lancet Respiratory Medicine. 2022 Apr;10(8).

[16] Townsend L, Dyer AH, Jones K, Dunne J, Mooney A, Gaffney F, et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. Madeddu G, editor. PLOS ONE. 2020 Nov 9;15(11):e0240784.

[17] Bai F, Tomasoni D, Falcinella C, Barbanotti D, Castoldi R, Mulè G, et al. Female gender is associated with “long COVID” syndrome: a prospective cohort study. Clinical

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Microbiology and Infection. 2021 Nov;28(4).

[18] Peghin M, Palese A, Venturini M, De Martino M, Gerussi V, Graziano E, et al. Post-COVID-19 symptoms 6 months after acute infection among hospitalized and non-hospitalized patients. Clinical Microbiology and Infection. 2021 Oct;27(10):1507–13.

[19] Kim Y, Kim S-W, Chang H-H, Kwon KT, Hwang S, Bae S. One Year Follow-Up of COVID-19 Related Symptoms and Patient Quality of Life: A Prospective Cohort Study. Yonsei Medical Journal. 2022;63(6):499.

[20] Abboud H, Abboud FZ, Kharbouch H, Arkha Y, El Abbadi N, El Ouahabi A. COVID-19 and SARS-Cov-2 Infection: Pathophysiology and Clinical Effects on the Nervous System. World Neurosurgery. 2020 Aug;140:49–53.

[21] Petersen MS, Kristiansen MF, Hanusson KD, Danielsen ME, á Steig B, Gaini S, et al. Long COVID in the Faroe Islands: A Longitudinal Study Among Nonhospitalized Patients. Clinical Infectious Diseases. 2020 Nov 30;73(11).

[22] Augustin M, Schommers P, Stecher M, Dewald F, Gieselmann L, Gruell H, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. The Lancet Regional Health – Europe [Internet]. 2021 Jul 1 [cited 2021 Nov 5];6. Available from: https://www.thelancet.com/journals/lanepe/article/PIIS2666-7762(21)00099-5/fulltext#coronavirus-linkback-header

[23] Seeßle J, Waterboer T, Hippchen T, Simon J, Kirchner M, Lim A, et al. Persistent symptoms in adult patients one year after COVID-19: a prospective cohort study. Clinical Infectious Diseases. 2021 Jul 5;74(7).

[24] Stewart S, Newson L, Briggs TA, Grammatopoulos D, Young L, Gill P. Long COVID risk - a signal to address sex hormones and women's health. The Lancet Regional Health -

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Europe. 2021 Nov;11:100242.

[25] Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a Delphi consensus. The Lancet Infectious Diseases. 2021 Dec 21;22(4).

[26] R Core Team. R: The R Project for Statistical Computing [Internet]. R-project.org. 2019. Available from: https://www.r-project.org/

[27] Biorender. BioRender [Internet]. Home. biorender-marketing-site; 2019. Available from: https://biorender.com

[28] Office for National Statistics. Statements and letters - Office for National Statistics [Internet]. www.ons.gov.uk. 2020 [cited 2022 Sep 17]. Available from: https://www.ons.gov.uk/news/statementsandletters/

[29] CDC. Nearly One in Five American Adults Who Have Had COVID-19 Still Have “Long COVID” [Internet]. www.cdc.gov. 2022 [cited 2022 Sep 17]. Available from: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/20220622.htm#:~:text=Overall%2C%201%20in%2013%20adults

[30] Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. Nature Medicine. 2022 Aug 1;28(8):1706–14.

TABLES

Table 1: Clinical, sociodemographic characteristics and frequency of residual symptoms for COVID-19 and non-COVID-19 cases.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Characteristic [1]	COVID-19 cases n = 412	Non-COVID-19 cases n = 402
Age	36 (28-48)	34(25-43)
Age group		
15-29	126 (30.6%)	167 (41.5%)
30-39	123 (29.9%)	106 (26.4%)
40-49	75 (18.2%)	82 (20.4%)
>50	88 (21.4%)	47 (11.7%)
Female sex	228 (55.3%)	259 (64.4%)
Body mass index (BMI)	27 (23.5-29.8)	26.1 (23.5-29.1)
Number of acute symptoms	11 (8-14)	10 (7-13)
Comorbidities	148 (35.9%)	131 (32.5%)
Diabetes mellitus	12 (11.0%)	11 (2.7%)
Respiratory allergy	95 (23.3%)	126 (31.3%)
Obesity	102 (24.8%)	86 (21.4%)
Normal	310 (75.2%)	316 (78.6%)
Obese (BMI 30-34)	21 (5.1%)	17 (4.2%)
Overweight (BMI 35-39)	79 (19.2%)	67 (16.7%)

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Extremely obese (BMI >40)	2 (0.5%)	2 (0.5%)
Number of comorbidities		
1	101 (24.5%)	90 (22.4%)
>1	47 (11.4%)	41 (10.2%)
Type of Hospitalization		
Outpatient (mild)	346 (84.2%)	395 (98.5%)
Hospitalization non-ICU (moderate)	59 (14.4%)	6 (1.5%)
Hospitalization ICU (severe)	6 (1.5%)	0 (0%)
At least one residual symptom > 1 month after disease onset	122 (29.6%)	24 (5.8%)
Number of residual symptoms (median)	1 (1-2)	2 (1-3)
1	65 (15.8%)	9 (2.2%)
>1	57 (13.2%)	13 (5.4%)
Call Time Persistent Symptoms	3.35 (2.1-5.2)	3.65 (2.1-4.9)

[1]n(%) or Median (IQR)

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: Clinical, sociodemographic characteristics and frequency of residual symptoms for COVID cases with or without sequelae.

Characteristic [1]	No Sequelae n=290	Sequelae n=122
Age	35 (28, 47)	40 (32, 51)
Age group		
15-29	99 (34.1%)	27 (22.1%)
30-39	85 (29.3%)	38 (31.1%)
40-49	51 (17.9%)	24 (19.7%)
>50	55 (19.0%)	33 (27.0%)
Female sex	151 (52.1%)	77 (63.1%)
BMI	26.9 (23.5, 29.7)	27.2 (23.8, 30.1)
Number of acute symptoms	11 (7, 14)	12 (9, 15)
Comorbidities	101 (34.8%)	47 (38.5%)
Diabetes mellitus	6 (2.1%)	6 (2.1%)
Respiratory allergy	63 (21.9%)	32 (26.2%)
Obesity	70 (24.1%)	32 (26.2%)
Obese (BMI 30-34)	12 (4.1%)	9 (7.4%)
Overweight (BMI 35-39)	57 (19.7%)	22 (18.0%)

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Extremely obese (BMI >40)	1 (0.3%)	1 (0.3%)
Number of comorbidities		
1	75 (25.9%)	26 (21.3%)
>1	26 (9.0%)	21 (17.2%)
Time of disease onset	3.30 (2.00, 5.10)	3.75 (2.12, 5.55)
Hospitalization	44 (15.2%)	21 (17.2%)

1. n(%) or Median (IQR)

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Multivariable analysis of odds ratio (lower and upper confidence intervals) for presenting long COVID symptoms.

	OR	LCI*	UCI*	P value
Age group (reference 14-29 years				
30-39	1.71	0.95	3.12	0.078
40-49	2.04	1.02	4.10	0.045
≥ 50	2.44	1.29	4.66	0.006
Female sex	1.55	0.99	2.44	0.055
> 5 acute symptoms	3.15	1.37	8.55	0.012

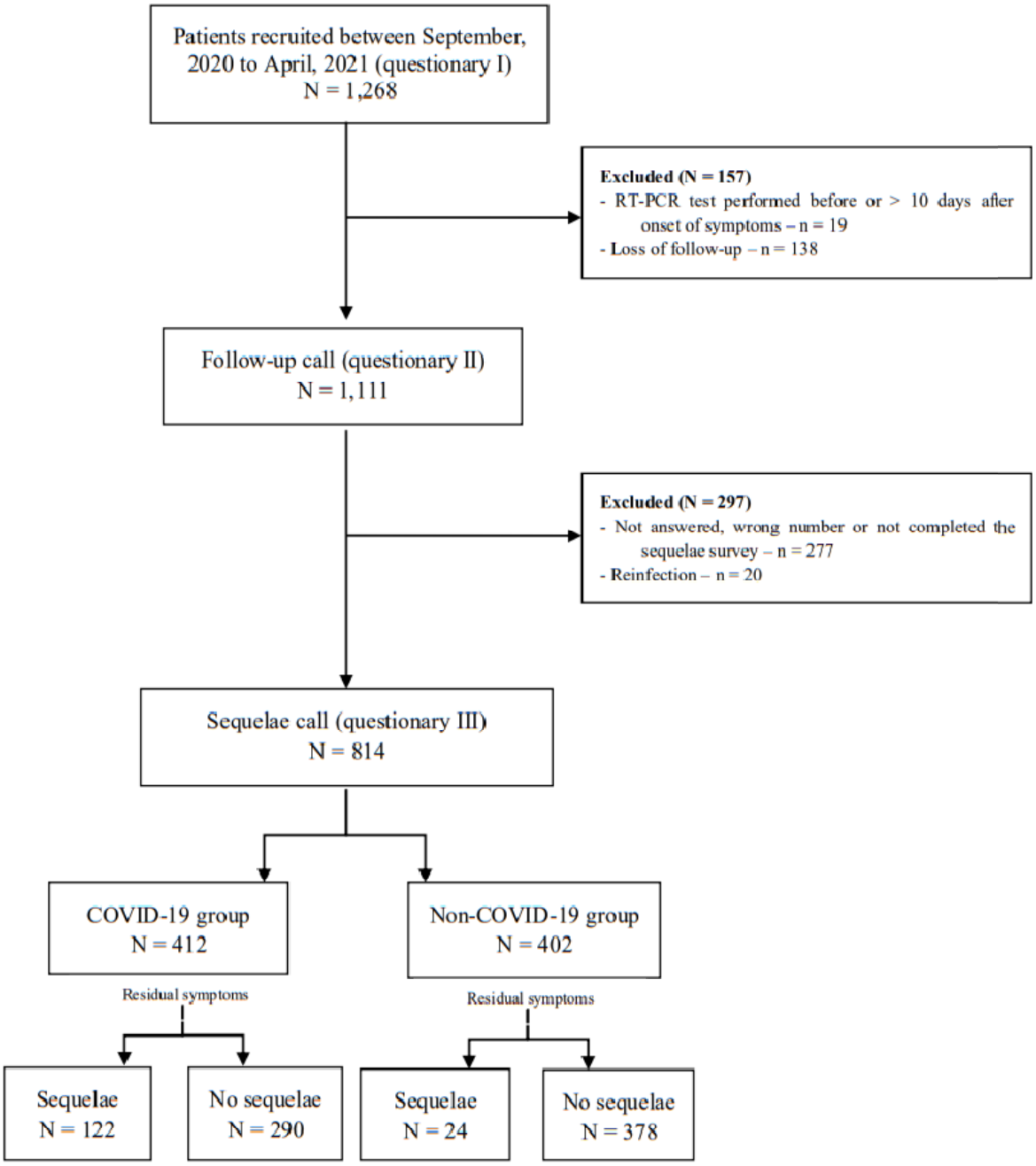
*LCI and UCI - lower and upper confidence intervals, respectively

FIGURES

medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

Fig 1. Flowchart of the study population.



medRxiv preprint doi: https://doi.org/10.1101/2022.09.15.22279958; this version posted September 22, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license .

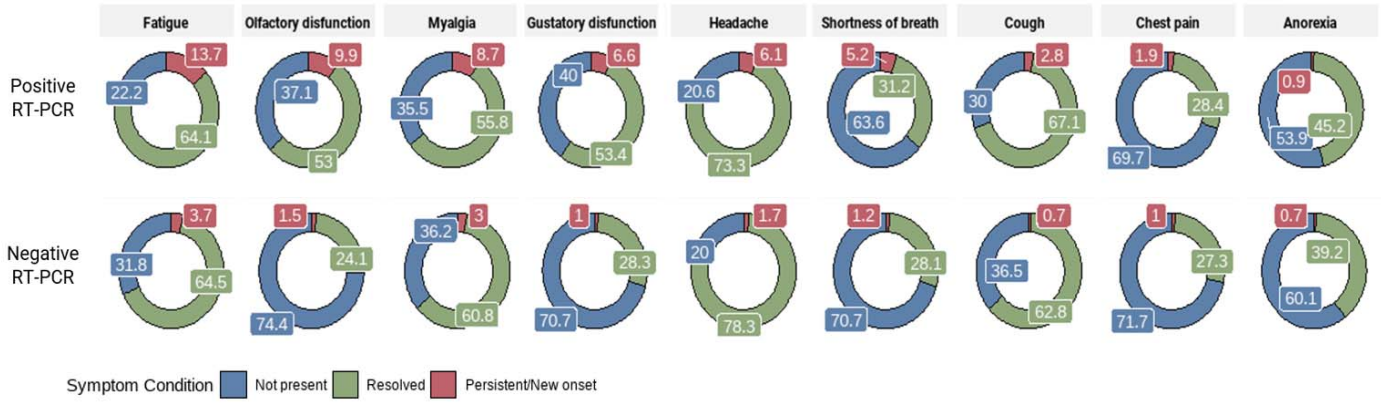



Fig. 2: Frequency of acute and residual symptoms for COVID-19 and non-COVID-19 cases.