medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Responsiveness to pulmonary rehabilitation in COPD is associated with changes in microbiota

Sara Melo-Dias[1,2,3], Msc, Miguel Cabral[1,3], Msc, Andreia Furtado[1,3], Msc, Sara Souto-Miranda[2,4], Msc, Maria Aurora Mendes[3,5], MD, João Cravo[5], MD, Catarina R. Almeida[1,3], PhD, Alda Marques[2,3], PhD‡ and Ana Sousa[1,3], PhD‡

Affiliations:
[1]Department of Medical Sciences, University of Aveiro, Aveiro, Portugal; [2]Lab3R – Respiratory Research and Rehabilitation Laboratory, School of Health Sciences (ESSUA), University of Aveiro, Aveiro, Portugal; [3]Institute of Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal; [4]Department of Respiratory Medicine, Maastricht University Medical Centre, NUTRIM School of Nutrition and Translational Research in Metabolism, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands.; [5]Department of Pulmonology of the Hospital Center of Baixo Vouga
‡These authors contributed equally to this study.

Correspondence: Ana Sousa, Department of Medical Sciences, Institute of Biomedicine, University of Aveiro, 3810-193 Aveiro, Portugal (amsousa@ua.pt)

Conflicts of interest: S. Melo-Dias has nothing to disclose. M. Cabral has nothing to disclose. A. Furtado has nothing to disclose. S. Souto-Miranda has nothing to disclose. M. A. Mendes has nothing to disclose. J. Cravo has nothing to disclose. C. R. Almeida has nothing to disclose. A. Marques has nothing to disclose. A. Sousa has nothing to disclose.

Funding: This study is integrated in "PRIME – Pulmonary Rehabilitation and microbiota in exacerbations of COPD", "GENIAL – Genetic and Clinical markers of COPD trajectory", and "MicroAgeing - The role of microbiota in ageing", funded by Programa Operacional de Competitividade e Internacionalização - COMPETE, through Fundo Europeu de Desenvolvimento Regional - FEDER (POCI-01-0145-FEDER-007628 and POCI-010145-FEDER-028806), Fundação para a Ciência e Tecnologia - FCT (PTDC/SAU-SER/28806/2017, PTDC/DTP-PIC/2284/2014 and PTDC/BIA-EVL/30212/2017) and under the project UIDB/04501/2020. S. Melo-Dias was supported by Grant SFRH/BD/140908/2018 from FCT. S. Souto-Miranda was supported by Grant SFRH/BD/146134/2019. A. Sousa was funded from national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the Scientific Employment Stimulus - Institutional Call - reference CEECINST/00026/2018.

Notation of prior abstract publication/presentation:
S. Melo-Dias, M. Cabral, A. Furtado, S. Souto-Miranda, J. Cravo, M. A. Mendes, C. R. Almeida, A. Marques, A. Sousa, 2022, **"Responsiveness to pulmonary rehabilitation is related with changes in oral microbiota of people with COPD"** (Oral communication in 38º Congresso de Pneumologia, November 2022)
S. Melo-Dias, M. Cabral, A. Furtado, S. Souto-Miranda, J. Cravo, M. A. Mendes, C. R. Almeida, A. Marques, A. Sousa, 2022, **"Pulmonary rehabilitation changes the oral microbiota of people with COPD"** (Oral communication in ERS International Congress, September 2022)
S. Melo-Dias, M. Cabral, A. Furtado, S. Souto-Miranda, C. R. Almeida, A. Marques, A. Sousa, 2022, **"The Effectiveness of pulmonary rehabilitation in COPD is associated with specific shifts in oral microbiota"** (Poster presentation in the 7th meeting of International Society for Evolution, Medicine and Public Health, July 2022)
A. Furtado, S. Melo-Dias, M. Cabral, A. Marques, A. Sousa, 2021, **"The effect of pulmonary rehabilitation in salivary microbiota of people with chronic obstructive pulmonary disease: A longitudinal study"** (Poster presentation in IV International Conference - "Microbiota and Health", 22 October 2021)

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/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Abstract

Background: Pulmonary Rehabilitation (PR) is one of the most cost-effective therapies for chronic obstructive pulmonary disease (COPD) management. There are, however, people who do not respond to PR and reasons for non-response are mostly unknown. PR is likely to change the airway microbiota and this could play a role in its responsiveness. In this study we have explored the association between PR effectiveness and specific alterations in oral microbiota and inflammation.

Methods: A prospective longitudinal study was conducted. Data on exercise capacity, dyspnoea, impact of disease and 418 saliva samples were collected from 76 patients, half of whom participated in a 12-weeks PR programme. Responders and non-responders to PR (dyspnoea, exercise-capacity and impact of disease) were defined based on minimal clinically important differences.

Results: Changes in microbiota, including *Prevotella melaninogenica* and *Streptococcus* were observed upon PR. *Prevotella*, previously found to be depleted in severe COPD, increased during the first month of PR in responders. This increase was negatively correlated with *Streptococcus* and *Lautropia*, known to be enriched in severe cases of COPD. Simultaneously, an anti-inflammatory commensal of the respiratory tract, *Rothia,* correlated strongly and negatively with several pro-inflammatory markers, whose levels were generally boosted by PR.

Conversely, in non-responders, the observed decline in *Prevotella* correlated negatively with *Streptococcus* and *Lautropia* whose fluctuations co-occurred with several pro-inflammatory markers.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Conclusions: PR is associated with changes in oral microbiota. Specifically, PR increases salivary *Prevotella melaninogenica* and avoids the decline in *Rothia* and the increase in *Streptococcus* and *Lautropia* in responders, which may contribute to the benefits of PR.

Key words: oral microbiota; inflammation; COPD; pulmonary rehabilitation; responsiveness

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Introduction

Chronic obstructive pulmonary disease (COPD) is heterogenous and complex and therefore, difficult to treat and manage. Pulmonary rehabilitation (PR) is a grade A non-pharmacological therapy and one of the most cost-effective approaches for the management of COPD [1]. Compared to pharmacological treatment, PR is three to five times more effective in improving exercise capacity, dyspnoea, and quality of life [2]. Response to PR is however multidimensional and heterogeneous, which means that for the same outcome patients are not equally responsive [3]. Reasons behind non-response are mostly unknown [3,4].

An association between the airway microbiota, including the oral microbiota [5], and disease severity has been extensively established in people with COPD [6–8]. However, the impact of PR on the airway microbiota has not yet been investigated, mainly because of the difficulty in collecting airway samples, such as sputum and bronchoalveolar lavage (BAL), routinely. Considering that only around 30% of people with COPD have productive cough [1], inducing sputum production becomes the only viable alternative. Unfortunately, this is challenging to collect outside the hospital context and requires both specialized personnel and equipment. The acquisition of BAL samples is even more challenging, as it is limited to hospital centres and bronchoscopy in patients with COPD carries a significantly higher risk of complications such as pneumonia, respiratory failure and desaturation compared with those with normal lung function [9]. In this scenario, saliva emerges as a viable non-invasive alternative to sampling the airway microbiota, which collection is simple and can be performed in different settings, namely at-home. Moreover, the oral and lower airway microbiotas are highly correlated [10], given the

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

topological continuity between the niches, implying oral bacteria as the major colonizers of the lungs through aspiration [10–12].

An additional reason to choose saliva to evaluate the effect of PR in the airway microbiota, is the role of oral bacteria in nitrate metabolism. Exercise training, one of the main components of PR, stimulates the synthesis of nitric oxide by the human body, which is a key regulator of skeletal muscle blood flow, contractility, and mitochondrial function [13]. Oral microbiota has been strongly implicated in exercise performance, mainly due to its essential role in nitrate-nitrite-nitric oxide pathway [14,15], with recent studies showing nitrate oral supplementation to enhance PR effectiveness [16]. The oral microbiota although acknowledged as essential for this positive effect, was never studied.

On the other hand, microbiota modifications are frequently connected with an inflammatory response, which is well known to be modulated by exercise [17] and affected by PR, although with inconsistent results [18–20].

Here, we have explored the association between PR and changes in oral microbiota and inflammatory markers of people with COPD to propose that PR effectiveness could be related, at least partially, with bacterial-driven immune regulation.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Methods

A prospective longitudinal cohort study was conducted. Ethical approvals were obtained from Administração Regional de Saúde Centro (64/2016), Centro Hospitalar do Baixo Vouga (including Estarreja's Hospital; 08-03-17) and Agrupamento dos Centros de Saúde do Baixo Vouga. Written informed consent was obtained from all participants.

The study was reported following the "Strengthening the Reporting of Observational Studies in Epidemiology" statement [21]. Participants were identified and referenced by clinicians who briefly explained the purposes of the study. Participants were eligible if diagnosed with COPD according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria [1] and stable with no acute exacerbations in the month prior to enrolment. Exclusion criteria were presence of severe cardiac, musculoskeletal, or neuromuscular diseases, signs of cognitive impairment, active neoplasia or immune diseases, that could hinder their participation in PR.

The opportunity to be integrated in the PR programme was offered to all participants. Those who did not accept to participate in the intervention but agreed with being monitored during the 5-month period were included in the control group. The intervention group (n=38) undertook a 12-week community-based PR program, whereas the control group (n=38) did not.

Sociodemographic (age, sex), anthropometric (height and weight) and general clinical (smoking habits; medication, long-term oxygen use; number of acute exacerbations and hospitalizations in the past year and comorbidities - Charlson Comorbidity Index) data were collected with a structured questionnaire. Lung function was assessed with spirometry as recommended. Exercise capacity was assessed with the six-minute walk

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

test (6MWT), a self-paced test of walking capacity. Impact of the disease was assessed with the COPD assessment test (CAT), a 8-item questionnaire, each assessed with a 6-point Likert scale. Dyspnoea at rest was assessed with the modified Borg Scale (mBorg), a 10-item scale. Clinical data (pre-post PR for the intervention group and M0-M3 for the control group) was collected with a structured protocol adapted from the team published work [22].

Saliva samples were collected monthly with the passive drool method. Prior to sample collection, patients were advised to drink a glass of water (especially if they had recently drunk coffee or citrus juice) and to provide 3-4 mL of saliva using a labelled sample collection cup. Subsequently, the sample was transported in a cooler to the lab as quickly as possible and preserved at -80ºC until DNA extraction and/or inflammatory markers' quantification (stored at -80ºC for 1 to 6 months).Response/non-response to PR was determined based on published minimal clinical importance differences (MCIDs), i.e., -1 point for mBorg [23]; 25m for 6MWT [24] and -2 points for CAT [25].

Deep sequencing of the V4 hypervariable region of 16S rRNA gene (F515/R806 primer pair) was performed for all samples. The bead assay LEGENDplex™ Human Inflammation Panel 1 (13-plex) (BioLegend, San Diego, CA, USA) was used to quantify inflammatory markers (IL-1β, IFN-α2, IFN-γ, TNF-α, MCP-1, IL-6, IL-8, IL-10, IL-12p70, IL-17A, IL-18, IL-23, and IL-33) in both groups in three timepoints, M0, M1 and M3. QIIME2 2020.8 [26,27] was used to perform microbiota analyses. All statistical analyses were performed in GraphPad Prism 8 [28] and R software v 3.6.1 [29]. Longitudinal models (*lme4* package [30]) and repeated-measures correlations (*rmcorr* package [31]) were performed in R software v 3.6.1 [29]. The full description of the methodology used,

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

including all steps of data collection, processing, and data analyses, are available in Supplementary file 1.

Results

Cohort characterisation

Seventy-six individuals with COPD were included in this study, 38 (29 male, 72±9y, BMI: 26±4kg/m^2, FEV_1pp 49.2±16 % predicted, GOLD A-8, B-20, C-0, D-10) in the intervention group and the remaining 38 patients (31 male, 70±8y, BMI: 26.4±4.8kg/m^2, FEV_1pp 52.3±19.8 % predicted, GOLD A-17, B-12, C-1, D-8) in the control group. Table 1 summarizes the baseline characteristics of the two groups. No significant differences were found between groups (see Table 1), and therefore, no adjustment for confounding factors was performed.

Table 1 – Sociodemographic, anthropometric, and clinical characteristics of individuals with chronic obstructive pulmonary disease at beginning of the study (n=76).

Characteristics	Intervention group (n=38)	Control group (n=38)	*p*-value
Age (years), mean±SD	72±9	70±7.60	0.19[1]
Gender, n (%)			0.57[2]
Male	29 (76%)	31 (82%)	
Female	9 (24%)	7 (18%)	
BMI (kg/m^2), mean±SD	26±4	26.4±4.8	0.72[3]
FEV_{1PP}, mean±SD	49.2±16.0	52.3±19.8	0.53[1]
FVC_{PP}, mean±SD	77.8±17.3	76.9±27.9	0.68[1]
Pack-years, mean±SD	42.4±43.0	53.0±51.2	0.27[1]
Comorbidities (CCI), mean±SD	4.2±1.4	3.8±1.3	0.10[1]
GOLD grade, n (%)			0.22[4]
1	3 (8%)	6 (16%)	
2	14 (37%)	12 (32%)	
3	20 (53%)	15 (39%)	
4	1 (3%)	5 (13%)	
GOLD group, n (%)			0.069[4]

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

A	8 (21%)	17 (45%)	
B	20 (53%)	12 (32%)	
C	0 (0%)	1 (3%)	
D	10 (26%)	8 (21%)	
Medication, n (%)			
ICS	21 (55%)	19 (50%)	0.8[4]
LABA	33 (87%)	30 (79%)	0.36[4]
LAMA	24 (63%)	25 (66%)	>0.99
LTRA	2 (5%)	2 (5%)	>0.99[4]
SABA	9 (24%)	5 (13%)	0.37[4]
SAMA	2 (5%)	0 (5%)	0.49[4]
Macrolides in stability	0 (0%)	0 (0%)	>0.99
Number of exacerbations in the last year, n (%)			0.60[2]
0 or 1	27 (71%)	29 (76%)	
≥2 or 1 with a hospitalisation	11 (29%)	9 (24%)	
Number of hospitalisations in the last year due to COPD, n (%)			0.67[4]
0	32 (84%)	34 (89%)	
1	2 (5%)	3 (8%)	
2	3 (8%)	1 (3%)	
Missing values	1 (3%)	0 (0%)	
mBORG, mean±SD	0.8±1.3	0.5±1.0	0.3[1]
CAT, mean±SD	17.1±8.1	11.7±7.4	0.003[1]
Walked distance in 6MWT, mean±SD	389.0±133.9	457.5±82.8	0.045[1]

SD: standard deviation, n (%): absolute and relative frequency, respectively; BMI: body mass index, FEV_{1pp}: forced expiratory volume in 1 second percent predicted, CCI: charlson comorbidity index, GOLD: global initiative for chronic obstructive lung disease, ICS: inhaled corticosteroids, LABA: long-acting beta agonists, LAMA: long-acting muscarinic agonists, LTRA: leukotriene receptor antagonists, SABA: short-acting beta agonists, SAMA: short-acting muscarinic agonists, [1] p-value obtained with Mann Whitney U-test, [2] p-value obtained with Chi-square test, [3] p-value obtained with unpaired t-test with Welch's correction, [4] p-value obtained with Fisher's exact test.

Oral microbiota was similar between groups, being composed by two major phyla, Firmicutes (~43%) and Bacteroidetes (~24%), followed by Proteobacteria, Fusobacteria, Actinobacteria and ten low abundant phyla (<2%) (Supplementary figure 1).

Changes in oral microbiota and inflammatory response are associated with PR

We followed the temporal dynamics of beta-diversity (global rate of change (grc) of Weighted-Unifrac distance) in patients for a period of five months, including the 12 weeks of PR, to understand the impact of PR in microbiota composition. As a control for

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

steady-state microbiota fluctuations, an independent cohort of patients who did not undergo PR, was surveyed for an equivalent amount of time.

The dynamics of microbiota composition was significantly different between patients undergoing PR and controls (Figure 1A) (lmer on Weighted-Unifrac distance (grc), Group:Time-point, F= 7.58, p<0.001).

Principal coordinate analysis of pairwise distances (Weighted Unifrac) between groups in each timepoint showed significant differences in microbiota composition between groups upon one month of PR (Figure 1B) (PERMANOVA, F=2.3, *p*=0.038). The top four amplicon sequence variants (ASVs) responsible for individual's separation after the first month of PR were identified as *Veillonella*, *Prevotella melaninogenica*, *Haemophilus parainfluenzae* and *Streptococcus.*

Regarding microbiota diversity within individuals (alpha diversity), no significant changes were observed upon PR (Figure 1C) (Wilcoxon test on phylogenetic diversity: W= -85. p=0.48; Wilcoxon test on Shannon diversity: W= -1. p>0.99).

Next, we compared the microbiota composition of the intervention group at baseline with that at M1, M2 and M3, to identify which taxa could have been affected by PR. PERMANOVA (M0vsM1: F=0.6, p_{adjust}>0.99; M0vsM2: F=1.3, p_{adjust}>0.99; M0vsM3: F=1.2, p_{adjust}>0.99; M0vsM4: F=1, p_{adjust}>0.99; M0vsM5: F=1.4, p_{adjust}=>0.99), ANCOM and LEfSe detected no significant differences between pre- and post-PR microbiota composition. This is compatible with the inherent correlation among the groups (composed by the same individuals) being stronger than differences introduced by PR.

To overcome this limitation, given the longitudinal design of the experiment, we next tested for differences in taxa dynamics between the intervention and control group. This

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

was performed using linear mixed-effect models. The temporal dynamics of the phylum Fusobacteria (lmer, Group:Time-point, F= 5.9, p=0.02) and several oral genera, such as *Fusobacterium*, *Streptococcus*, *Dialister* and *Selenomonas* were observed to be associated with PR (Table 2). Target testing for the top four ASVs distinguishing patients on M1 was also performed. Remarkably, significant alterations in the relative abundances of *Prevotella melaninogenica* and *Streptococcus*, two ASVs previously associated with disease severity in people with COPD [5], were also found (lmer, *P. melaninogenica:* Group:Time-point, F= 4.34, p=0.04; lmer, *Streprococcus:* Group:Time-point, F= 4.96, p=0.03).

Table 2 – Summary table of significant linear mixed effect models established to assess differences in microbiota dynamics between intervention and control groups.

OTUs/ASVs longitudinal dynamics	test's statistics	*p*-value (TP:Group)
Fusobacteria	5.86	0.016
Atopobium	5.51	0.019
Streptococcus	5.03	0.026
Dialister	6.39	0.012
Parvimonas	4.75	0.030
Fusobacterium	7.00	0.009
a5189f77a2cfeab3bc1602ff5c8ac3e9: Streptococcus	4.96	0.027
d0b698c7298bf04110a6d2f220879bfb: Prevotella melaninogenica	4.34	0.038
e27680d4009f98f30248d823bc17fb8e: Haemophilus parainfluenzae	0.18	0.673
5608c3e6c9de9ceb79610e7786bd0ac4: Veillonella	3.30	0.070

Models were produced based on the longitudinal relative frequencies of phyla, genera, and ASVs *a5189f77a2cfeab3bc1602ff5c8ac3e9, d0b698c7298bf04110a6d2f220879bfb, e27680d4009f98f30248d823bc17fb8e* and *5608c3e6c9de9ceb79610e7786bd0ac4*. Data upon relative frequencies was transformed with arcsine and square root transformation. TP: time-point

Considering that changes in microbiota are expected to be related with the inflammatory response, we further quantified a panel of 13 cytokines in three time points: M0, M1 and M3 in both intervention and control groups (Figure 2 and

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Supplementary figure 2). Significant changes in the inflammatory profile of patients undergoing PR were observed (Figure 2). Specifically, upon one month of PR (M1) significant increases in the amounts of IL-10 and IL-18 were observed (Wilcoxon signed-rank: IL-10, W=163 p_{adjust}=0.038; IL-18, W=165 p_{adjust}=0.035). No significant changes were observed in the same period in control group. After three months of PR (M3), TNF-α, IL-1β, IL-18, IL-10 and IL-6 were also significantly higher than at baseline (M0) (Wilcoxon signed-rank: IL-1β, W=205 p_{adjust}=0.016; TNF-α, W=251 p_{adjust}=0.002; IL-10, W=263 p_{adjust}=0.0008; IL-6, W=189 p_{adjust}=0.03; IL-18, W=233 p_{adjust}=0.004). IL-10 and IL-6 also increased in the control group over the same time interval (Wilcoxon signed-rank: IL-10, W=225 p_{adjust}=0.018; IL-6, W=224 p_{adjust}=0.019), suggesting that these cytokines may be less stable.

Regarding comparisons between the groups, shifts in IL-10 and IL-18 were significantly higher after 1 month of PR than the equivalent shifts in the control group (Mann-Whitney U-test, IL-10, U=180 p=0.043; IL-18, U=173 p=0.03). Over 3 months of PR, TNF-α, IL-18 and IL-10 had significantly higher shifts in the intervention than in the control group (Mann-Whitney U-test, TNF-α, U=246 p=0.04; IL-10, U=196 p=0.0004; IL-18, U=249 p=0.046). No additional significant differences were found in both groups (Supplementary figure 2).

Interestingly, the variance associated with cytokine shifts over time was generally higher in the intervention than in the control group (Supplementary table 2). This is compatible with PR generally influencing cytokine levels.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Oral microbiota and inflammatory response vary differentially among responders and non-responders during PR

PR improved all outcomes significantly (Wilcoxon test, $p < 0.05$), except for BMI (W=-114, p=0.31) and mBorg (W=37, p=0.5) (Supplementary table 3). Mean improvements only exceeded the MCID for exercise capacity (6MWT, mean-diff: 45.43; Cohen's d=0.34) (Supplementary table 2). Supplementary figure 3 represents the overlap between responders (R) and non-responders (NR) to PR in exercise capacity (6MWT), impact of disease (CAT) and dyspnoea (mBorg). From 38 patients, 24% responded in dyspnoea, 63% responded in exercise capacity and 63% responded in the impact of disease. 16% of patients responded simultaneously to all three domains and 16% failed to respond to any domain.

We analysed differences in microbiota composition among R and NR at M0 to query for a possible relationship between microbiota composition before PR and its effectiveness (Figure 3A). Principal coordinate analysis of pairwise distances (Weighted-Unifrac) showed significant differences between R and NR to dyspnoea (PERMANOVA $p = 0.04$) and captured 66% of total diversity (Figure 3B). Furthermore, the ASV responsible for the biggest separation between individuals in this analysis, was *P. melaninogenica*, whose frequency was below average in 75% of R. In accordance, the differential abundance analysis (LEfSe) showed that *Prevotella* (Figure 3C) was significantly enriched in NR to dyspnoea, with an effect-size of 9. No significant differences were observed for the other domains.

Next, we used linear mixed effect models to compare the microbiota temporal dynamics between R and NR. Despite being differentially represented in R and NR (to dyspnoea) before PR, the frequency trajectory of *P. melaninogenica* was not significantly different

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

among the two sets of patients during the intervention. Instead, several other taxa were found to have significantly different trajectories between the groups (Table 3 and Supplementary figure 4). Regarding the exercise capacity, the frequency trajectory of *Lautropia* (family *Burkholderiaceae*) was significantly different between R and NR (lmer, Group:Time-point, F= 2.9, p=0.02), reaching higher values in NR by the end of PR (Supplementary figure 3). The dynamics of *P. melaninogenica* was distinct among R and NR (lmer, Group:Time-point, F= 2.6, p=0.03) to impact of disease (Supplementary figure 4).

Table 3 – Summary table of significant linear mixed effect models established to assess differences in microbiota dynamics between responders and non-responders to dyspnoea, exercise capacity and impact of the disease.

OTUs/ASVs longitudinal dynamics	test's statistics	*p*-value (TP:Group)
Response to dyspnoea during exercise (Borg Modified Scale of Dyspnoea) MCID: -1 point		
Spirochaetes	2.33	0.046
TM7	5.40	0.000
Oribacterium	2.68	0.024
Dialister	2.69	0.023
Schwartzia	2.55	0.036
[Mogibacteriaceae]	2.45	0.036
Bulleidia	3.08	0.011
Treponema	2.33	0.046
Response to exercise capacity (6-minute walk test) MCID: +25m		
Lautropia	2.87	0.017
Response to impact of the disease (COPD assessment test) MCID: -2 points		
d0b698c7298bf04110a6d2f220879bfb: Prevotella melaninogenica	2.58	0.029

Models were produced based on the longitudinal relative frequencies of phyla, genera, *a5189f77a2cfeab3bc1602ff5c8ac3e9* and *d0b698c7298bf04110a6d2f220879bfb*. Data upon relative frequencies was transformed with arcsine and square root transformation. TP: time-point

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

We compared the salivary levels of cytokines present in R and NR to question whether responsiveness to PR was related with the inflammatory response. Distinct patterns of inflammatory response were observed between R and NR, in all domains (Supplementary figure 5). Significant shifts during PR were always towards values higher than the baseline and occurred mainly in NR patients. Specifically, after PR, levels of TNFα, IL-1β, IL-18 and IL-10 were higher in NR of all domains, whereas only TNF-α and IL-10 showed significant changes in R patients (please see Supplementary figure 5 for details on the statistical tests results). Interestingly, the variance in cytokine shifts after PR was also mostly higher in NR than in R (Supplementary tables 4–6).

PR effectiveness is related with specific bacteria-inflammation correlation signatures

Independent analyses revealed an association between PR effectiveness and either changes in oral microbiota or inflammatory markers. We explored the patterns of longitudinal correlation between the two to understand in what extent PR effectiveness is linked with specific bacteria-inflammation associations. Figure 4 represents the network of significant correlations between inflammatory markers and bacterial genera inferred for the three domains. Remarkably, R and NR presented distinct patterns of bacteria-cytokine correlation (Figure 4).

In all groups of NR, *Lautropia* showed a strong positive correlation with several pro-inflammatory cytokines, being the strongest with TNF-α in NR to exercise capacity (rmcorr = 0.65, p=0.002) and impact of the disease (rmcorr = 0.63, p=0.002). Conversely, *Rothia* presented a negative correlation with several pro-inflammatory markers in all groups of R, with the strongest correlation with TNF-α (rmcorr = -0.73, p=0.007). Furthermore, *Gemellaceae*, was also negatively correlated with several inflammatory

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

markers (strongest with IL-8, rmcorr = -0.69, $p<0.0001$) in R to exercise capacity and impact of the disease. Strikingly, except for a positive correlation between *Haemophilus* and IFN-α2 in R to impact of the disease (rmcorr = 0.43, $p=0.03$), all significant correlations with Proteobacteria were found in NR.

We then performed a longitudinal correlation analysis between bacterial frequencies to gain further insight on bacterial interactions that might be related with PR responsiveness but are not necessarily linked with the immune response (Figure 5). In an effort to be conservative, since compositional data are intrinsically correlated, we chose a subset of taxa for this analysis. We included the major hubs found in the correlation between bacteria and inflammatory markers (*Lautropia*, *Rothia*, *Gemellaceae* and *Kingella*), plus five other oral taxa previously associated [5] with severity (*Prevotella*, *P. melaninogenica*, *Streptococcus*, *Streptococcus* sp., *Haemophilus* and *Porphyromonas*).

In this analysis, *Prevotella*, whose frequency drops during PR in all NR (Supplementary figure 6 A-C), was negatively correlated with *Haemophilus* (rmcorr = -0.46, $p=0.005$), *Streptococcus* (rmcorr = -0.37, $p=0.03$) and *Lautropia* (rmcorr = -0.41, $p=0.01$) in NR to dyspnoea and with *Haemophilus* (rmcorr = -0.73, $p=0.0004$) in NR to exercise capacity. Conversely, during the first month of PR, the frequency of *P. melaninogenica* increased in all R (Supplementary figure 6A-C) and was negatively correlated with *Kingella* (R to dyspnoea, rmcorr = -0.59, $p=0.04$), *Lautropia* (R to exercise capacity, rmcorr = -0.52, $p=0.004$; R to impact of the disease, rmcorr = -0.43, $p=0.03$) and *Streptococcus* (R to impact of the disease, rmcorr = -0.41, $p=0.04$). *Gemellaceae* and *Rothia* were strongly positively correlated in all R (strongest in dyspnoea, rmcorr = 0.9, $p<0.0001$).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Discussion

We show, for the first time, that oral microbiota changes with PR and that shifts in specific taxa appear to be strongly related with its effectiveness. By tracking two independent groups of patients for almost half a year we were able to distinguish between steady-state microbiota fluctuations and those possibly induced by PR. Contrary to the simple expectation that a well-established therapy should positively impact on patients' microbiota, a more complex reality was unravelled. Overall, PR was not able to boost microbiota diversity within patients, a hallmark of dysbiosis, causing only low magnitude changes in a few taxa, and elicited the secretion of several pro-inflammatory cytokines.

However, when considering separately R and NR, taxa potentially related with responsiveness emerges and a differential inflammatory response becomes apparent. Previously, we have observed that low frequency of *Prevotella*, and in particular *P. melaninogenica*, was associated with a recent history of severe exacerbation [5]. Here we observed that in the first month of PR *Prevotella* was boosted in R of all domains. We hypothesize that this boost could be directly or indirectly related with the ability to respond. For example, by directly inducing the reduction of lung epithelial cell permeability (by modulating expression of tight junction proteins) [32], or indirectly by promoting lung innate immune responses, particularly against *Streptococcus pneumoniae* [33] and *Lautropia,* to which it was negatively correlated.

In fact, it has been recently observed that *P. melaninogenica* is able to promote *S. pneumoniae* rapid clearance from the lung via TLR2 activation, recruitment of

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

neutrophils and upregulation of TNFα and IL-10 [34]. Interestingly, it was also observed that the aspiration of a mixture of 3 species of human oral commensals in mice, including *P. melaninogenica*, decreased host´s susceptibility to *S. pneumoniae* [33].

Supporting the hypothesis that, by counteracting *Lautropia*, *Prevotella* could be positively associated with PR effectiveness, is the observation that *Lautropia* is enriched in the oral microbiota of more severe cases of COPD [5] and co-occur with high levels of inflammatory markers in the eosinophilic COPD [35], an endotype that has been suggested to be predictive of worse outcome upon PR [36] .

Conversely, in NR to dyspnoea and exercise capacity, the opposite scenario was observed, that is, a clear decline in *Prevotella* upon PR, negatively correlated with *Haemophilus* (exercise capacity) and with *Streptococcus* and *Lautropia* (dyspnoea). A pro-inflammatory role in COPD has been repeatedly attributed to the first two taxa [37] which could likely contribute to the lack of response to this therapy. Besides this, *Lautropia* positively correlated with several pro-inflammatory markers in all groups of NR.

Another striking observation, possibly related with the success of PR across the three domains, was the co-occurrence of *Rothia* and *Gemellacea* (in R) and its negative correlation with several pro-inflammatory markers during PR. *Rothia mucilaginosa* (the most abundant species of *Rothia* genus in our dataset) has been identified as an anti-inflammatory bacterium when present in the lungs (even in low abundance) of people with chronic respiratory diseases [38]. This has been attributed to its ability to reduce the levels of several pro-inflammatory cytokines and NF-kB activation in response to

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

immune stimulation by pathogenic bacteria like *Pseudomonas aeruginosa* and *Staphylococcus aureus* [38].

Considering the impact of PR on individual species, this was significant on the longitudinal dynamics of two particular ASVs, *P. melaninogenica* and *Streptococcus*, previously suggested to be associated with COPD severity [5]. Moreover, the longitudinal dynamics of *P. melaninogenica* was distinct among R and NR to impact of the disease, supporting its importance to PR effectiveness.

Besides its immunomodulatory properties, *P. melaninogenica* could have an additional role in PR effectiveness, since it belongs to the most prevalent genus of oral nitrate-reducing bacteria [39]. These are essential for the nitrate-nitrite-nitric oxide pathway which is implicated in exercise performance and recovery [14,15]. Coherently, it has been shown that nitrate oral supplementation increased the beneficial effects of PR [16].

Several inconsistencies have been reported regarding the effects of PR in the inflammatory status of people with COPD. Although it is unlikely that PR increases inflammation in the long-term, our findings corroborate the previous reported increase of TNF-α in plasma of people with COPD upon PR [23]. The observed increase of IL-10 suggest an attempt to counterbalance the rise of pro-inflammatory cytokines during PR, possibly via inhibition of nuclear factor kappa B (NF-kB) and the synthesis of pro-inflammatory cytokines [24].

Furthermore, it has been observed that an inflammatory response (measured by an increase in serum levels of IL-1β, TNF-α and IL-10) is necessary for recovery from

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

exercise-induced muscle damage [27, 28]. Given the overlap between inflammatory profiles of saliva and serum [25, 26], our results could reflect a similar process.

Some limitations of our study need to be acknowledged. First, the use of saliva for COPD assessment is still exploratory. Although the microbiota of upper and lower airways is highly correlated, with oral bacteria being the major colonizers of the lungs through microaspiration [10–12], more studies, besides Melo-Dias et al [5] validating its biological relevance in the context of COPD are needed. Additionally, despite longitudinal studies including PR being very resource demanding, further works with similar design should be carried out, ideally including multicentric trials with larger cohorts to evaluate the robustness of the findings. This is particularly important to assess the association between responsiveness and microbiota modulation. Second, clinical data collection was performed pre-post intervention and therefore it is not possible to adjust longitudinal models for multiple confounders such as occurrence of exacerbations, treatment with inhaled-corticosteroids and/or antibiotics.

Overall, besides responsiveness to PR being multidimensional and heterogeneous, giving rise to a moderate overlap in individuals response across domains, PR-induced changes in microbiota revealed surprisingly consistent patterns among R and NR. Furthermore, our findings suggest that PR effectiveness could be associated with a controlled inflammatory response to exercise, *i.e.*, the inflammatory response to exercise occurs accompanied by efficient regulatory mechanisms. These mechanisms can be mediated by bacteria and could be tested *in vitro*.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Future studies should address the implications and stability of these findings, clarifying the role of oral microbiota both as a biomarker of PR responsiveness and as a therapeutic target.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Acknowledgements:

Author contributions:

S. Melo-Dias and A. Sousa had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. A. Marques was responsible for ethical approval. A. Marques and A. Sousa were responsible for obtaining the funding, conceiving and designing the study. M. A. Mendes, J. Cravo and S. Souto-Miranda provided a substantial contribution for clinical data acquisition. S. Melo-Dias A. Furtado conducted sample processing, DNA extraction procedures and inflammatory markers quantification. S. Melo-Dias, M. Cabral, A. Furtado, C. R. Almeida and A. Sousa conducted bioinformatics and statistical analyses. S. Melo-Dias and A. Sousa drafted the manuscript. All authors critically revised the manuscript, ensured accuracy and integrity of the work and approved the final version to be published.

Availability of data and materials:

The dataset supporting the conclusions of this article is included within the article (and its additional file(s)). Furthermore, raw sequencing data was deposited in National Centre for Biotechnology Information's (NCBI) Sequence Read Archive (SRA) (BioProject PRJNA872131). Scripts for data analyses are provided in supplementary file 1.

Ethics approval and consent to participate:

A cross-sectional study was conducted. Ethical approvals were obtained from Administração Regional de Saúde Centro (64/2016) and from Centro Hospitalar do Baixo Vouga (08-03-17). Written informed consent was obtained from all participants.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Consent for publication:

Not applicable.

Acknowledgements:

We are grateful to all people with COPD who accepted to participate in this study and to the team of Respiratory Research and Rehabilitation Laboratory, School of Health Sciences (Lab3R-ESSUA) for their support.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

References

1 2022 GOLD Reports. Global Initiative for Chronic Obstructive Lung Disease - GOLD. https://goldcopd.org/2022-gold-reports-2/ (accessed 7 Jan 2022).

2 Casaburi R. Pulmonary Rehabilitation: Where We've Succeeded and Where We've Failed. *COPD: Journal of Chronic Obstructive Pulmonary Disease* 2018;**15**:219–22. doi:10.1080/15412555.2018.1503245

3 Augustin IML, Franssen FME, Houben-Wilke S, *et al.* Multidimensional outcome assessment of pulmonary rehabilitation in traits-based clusters of COPD patients. *PLOS ONE* 2022;**17**:e0263657. doi:10.1371/journal.pone.0263657

4 Spruit MA, Augustin IML, Vanfleteren LE, *et al.* Differential response to pulmonary rehabilitation in COPD: multidimensional profiling. *Eur Respir J* 2015;**46**:1625–35. doi:10.1183/13993003.00350-2015

5 Melo-Dias S, Valente C, Andrade L, *et al.* Saliva as a non-invasive specimen for COPD assessment. *Respiratory Research* 2022;**23**:16. doi:10.1186/s12931-022-01935-9

6 Budden KF, Shukla SD, Rehman SF, *et al.* Functional effects of the microbiota in chronic respiratory disease. *The Lancet Respiratory Medicine* 2019;**7**:907–20. doi:10.1016/S2213-2600(18)30510-1

7 Marsland BJ, Gollwitzer ES. Host–microorganism interactions in lung diseases. *Nat Rev Immunol* 2014;**14**:827–35. doi:10.1038/nri3769

8 Ubags NDJ, Marsland BJ. Mechanistic insight into the function of the microbiome in lung diseases. *European Respiratory Journal* 2017;**50**. doi:10.1183/13993003.02467-2016

9 Rand IAD, Blaikley J, Booton R, *et al.* British Thoracic Society guideline for diagnostic flexible bronchoscopy in adults: accredited by NICE. *Thorax* 2013;**68**:i1–44. doi:10.1136/thoraxjnl-2013-203618

10 Gaeckle NT, Pragman AA, Pendleton KM, *et al.* The Oral-Lung Axis: The Impact of Oral Health on Lung Health. *Respir Care* 2020;**65**:1211–20. doi:10.4187/respcare.07332

11 Mammen MJ, Scannapieco FA, Sethi S. Oral-lung microbiome interactions in lung diseases. *Periodontology 2000* 2020;**83**:234–41. doi:10.1111/prd.12301

12 Huffnagle GB, Dickson RP, Lukacs NW. The respiratory tract microbiome and lung inflammation: a two-way street. *Mucosal Immunol* 2017;**10**:299–306. doi:10.1038/mi.2016.108

13 Dyakova EY, Kapilevich LV, Shylko VG, *et al.* Physical exercise associated with NO production: signaling pathways and significance in health and disease. *Front Cell Dev Biol* 2015;**3**:19. doi:10.3389/fcell.2015.00019

14 Bescos R, Brookes ZLS, Belfield LA, *et al.* Modulation of oral microbiota: A new frontier in exercise supplementation. *PharmaNutrition* 2020;**14**:100230. doi:10.1016/j.phanu.2020.100230

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

15 Cutler C, Kiernan M, Willis JR, *et al.* Post-exercise hypotension and skeletal muscle oxygenation is regulated by nitrate-reducing activity of oral bacteria. *Free Radical Biology and Medicine* 2019;**143**:252–9. doi:10.1016/j.freeradbiomed.2019.07.035

16 Pavitt MJ, Tanner RJ, Lewis A, *et al.* Oral nitrate supplementation to enhance pulmonary rehabilitation in COPD: ON-EPIC a multicentre, double-blind, placebo-controlled, randomised parallel group study. *Thorax* 2020;**75**:547–55. doi:10.1136/thoraxjnl-2019-214278

17 Clauss M, Gérard P, Mosca A, *et al.* Interplay Between Exercise and Gut Microbiome in the Context of Human Health and Performance. *Frontiers in Nutrition* 2021;**8**.https://www.frontiersin.org/article/10.3389/fnut.2021.637010 (accessed 12 May 2022).

18 Gammal AIE, O'Farrell R, O'Mahony L, *et al.* Systemic Inflammatory Markers and Disease Severity in Chronic Obstructive Pulmonary Disease—The Effect of Acute Exercise and Pulmonary Rehabilitation. *Open Journal of Respiratory Diseases* 2015;**5**:19–31. doi:10.4236/ojrd.2015.51003

19 van der Vlist J, Janssen TWJ. The potential anti-inflammatory effect of exercise in chronic obstructive pulmonary disease. *Respiration* 2010;**79**:160–74. doi:10.1159/000233264

20 Gigliotti F, Coli C, Lombardelli L, *et al.* Effect of pulmonary rehabilitation (PR) on inflammation status in patients with COPD. *European Respiratory Journal* 2015;**46**. doi:10.1183/13993003.congress-2015.PA2210

21 Elm E von, Altman DG, Egger M, *et al.* Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8. doi:10.1136/bmj.39335.541782.AD

22 Marques A, Jácome C, Rebelo P, *et al.* Improving access to community-based pulmonary rehabilitation: 3R protocol for real-world settings with cost-benefit analysis. *BMC Public Health* 2019;**19**:676. doi:10.1186/s12889-019-7045-1

23 Jones PW, Beeh KM, Chapman KR, *et al.* Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014;**189**:250–5. doi:10.1164/rccm.201310-1863PP

24 Holland AE, Hill CJ, Rasekaba T, *et al.* Updating the minimal important difference for six-minute walk distance in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil* 2010;**91**:221–5. doi:10.1016/j.apmr.2009.10.017

25 Kon SSC, Canavan JL, Jones SE, *et al.* Minimum clinically important difference for the COPD Assessment Test: a prospective analysis. *Lancet Respir Med* 2014;**2**:195–203. doi:10.1016/S2213-2600(14)70001-3

26 Bolyen E, Rideout JR, Dillon MR, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7. doi:10.1038/s41587-019-0209-9

27 QIIME 2. https://qiime2.org/ (accessed 12 Jan 2022).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

28 Prism - GraphPad. https://www.graphpad.com/scientific-software/prism/ (accessed 12 Jan 2022).

29 R: The R Project for Statistical Computing. https://www.r-project.org/ (accessed 12 Jan 2022).

30 Bates D, Mächler M, Bolker B, *et al.* Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 2015;**67**:1–48. doi:10.18637/jss.v067.i01

31 Bakdash JZ, Marusich LR. Repeated Measures Correlation. *Frontiers in Psychology* 2017;**8**.https://www.frontiersin.org/article/10.3389/fpsyg.2017.00456 (accessed 15 Mar 2022).

32 Ramsheh MY, Haldar K, Esteve-Codina A, *et al.* Lung microbiome composition and bronchial epithelial gene expression in patients with COPD versus healthy individuals: a bacterial 16S rRNA gene sequencing and host transcriptomic analysis. *The Lancet Microbe* 2021;**2**:e300–10. doi:10.1016/S2666-5247(21)00035-5

33 Wu BG, Sulaiman I, Tsay J-CJ, *et al.* Episodic Aspiration with Oral Commensals Induces a MyD88-dependent, Pulmonary T-Helper Cell Type 17 Response that Mitigates Susceptibility to Streptococcus pneumoniae. *Am J Respir Crit Care Med* 2021;**203**:1099–111. doi:10.1164/rccm.202005-1596OC

34 Horn KJ, Schopper MA, Drigot ZG, *et al.* Airway Prevotella promote TLR2-dependent neutrophil activation and rapid clearance of Streptococcus pneumoniae from the lung. *Nat Commun* 2022;**13**:3321. doi:10.1038/s41467-022-31074-0

35 Wang Z, Locantore N, Haldar K, *et al.* Inflammatory Endotype–associated Airway Microbiome in Chronic Obstructive Pulmonary Disease Clinical Stability and Exacerbations: A Multicohort Longitudinal Analysis. *Am J Respir Crit Care Med* 2021;**203**:1488–502. doi:10.1164/rccm.202009-3448OC

36 Blood Eosinophils and Pulmonary Rehabilitation in COPD - PubMed. https://pubmed.ncbi.nlm.nih.gov/34777651/ (accessed 24 Jul 2022).

37 Celli BR, Wedzicha JA. Update on Clinical Aspects of Chronic Obstructive Pulmonary Disease. *N Engl J Med* 2019;**381**:1257–66. doi:10.1056/NEJMra1900500

38 Rothia mucilaginosa is an anti-inflammatory bacterium in the respiratory tract of patients with chronic lung disease | European Respiratory Society. https://erj.ersjournals.com/content/early/2021/09/16/13993003.01293-2021 (accessed 12 May 2022).

39 Burleigh MC, Liddle L, Monaghan C, *et al.* Salivary nitrite production is elevated in individuals with a higher abundance of oral nitrate-reducing bacteria. *Free Radic Biol Med* 2018;**120**:80–8. doi:10.1016/j.freeradbiomed.2018.03.023

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Figures

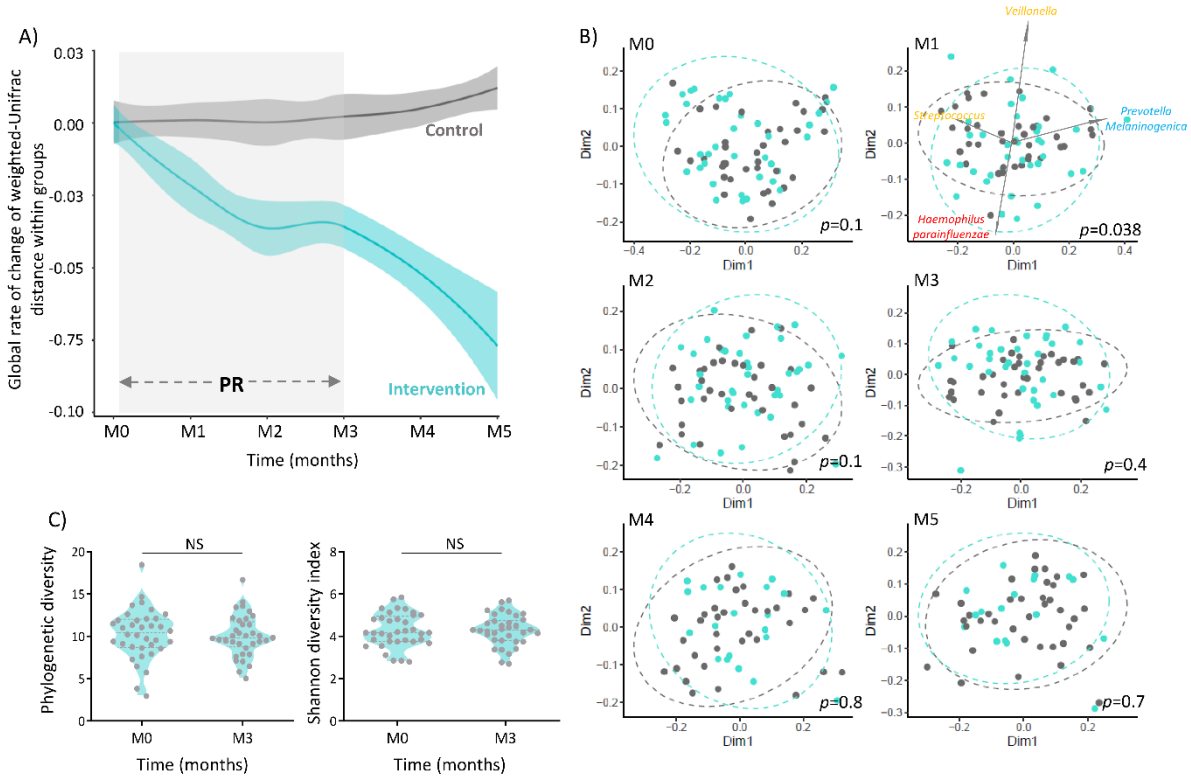


Figure 1. Pulmonary rehabilitation modulates the microbiota composition of people with COPD, but does not affect alpha diversity. A) Global rate of change of beta diversity was estimated by subtraction of baseline values at all timepoints. Blue and grey loess lines were fitted over the data points representing beta-diversity relative to baseline along time in the intervention and control groups, respectively. The blue and grey areas represent the 0.95 confidence intervals. (M0=immediately prior PR; M1, M2, M3= 1,2,3 months after initiating PR; M4, M5=1 and 2 months after finishing PR). Intervention and control group presented different microbiota dynamics over the 5-month follow-up (LMER on Weighted Unifrac distance matrix, Group:Time-Point, p<0.001). The grey rectangle indicates the time interval spanning PR. **B)** Principal coordinate analysis of Weighted Unifrac distance between groups in each timepoint. (M0, M1, M2, M3, M4 and M5). Significant differences in microbiota composition between groups were observed in M1 (PERMANOVA, F=2.27, *p*=0.038). Blue and grey dots and ellipses represent intervention and control groups, respectively. The biplot represented in M1 show the top ASVs contributing for sample segregation (*5608c3e6c9de9ceb79610e7786bd0ac4:*

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Veillonella, d0b698c7298bf04110a6d2f220879bfb: Prevotella melaninogenica, e27680d4009f98f30248d823bc17fb8e: Haemophilus parainfluenzae, a5189f77a2cfeab3bc1602ff5c8ac3e9: Streptococcus). **C)** No pre-post differences were observed in alpha diversity (Phylogenetic index and Shannon diversity index) in the Intervention group (Wilcoxon test: phylogenetic diversity: W= -85. p=0.48; Shannon diversity index: W=-1 p>0.99). **p*<0.05, ***p*<0.01, ****p*<0.001

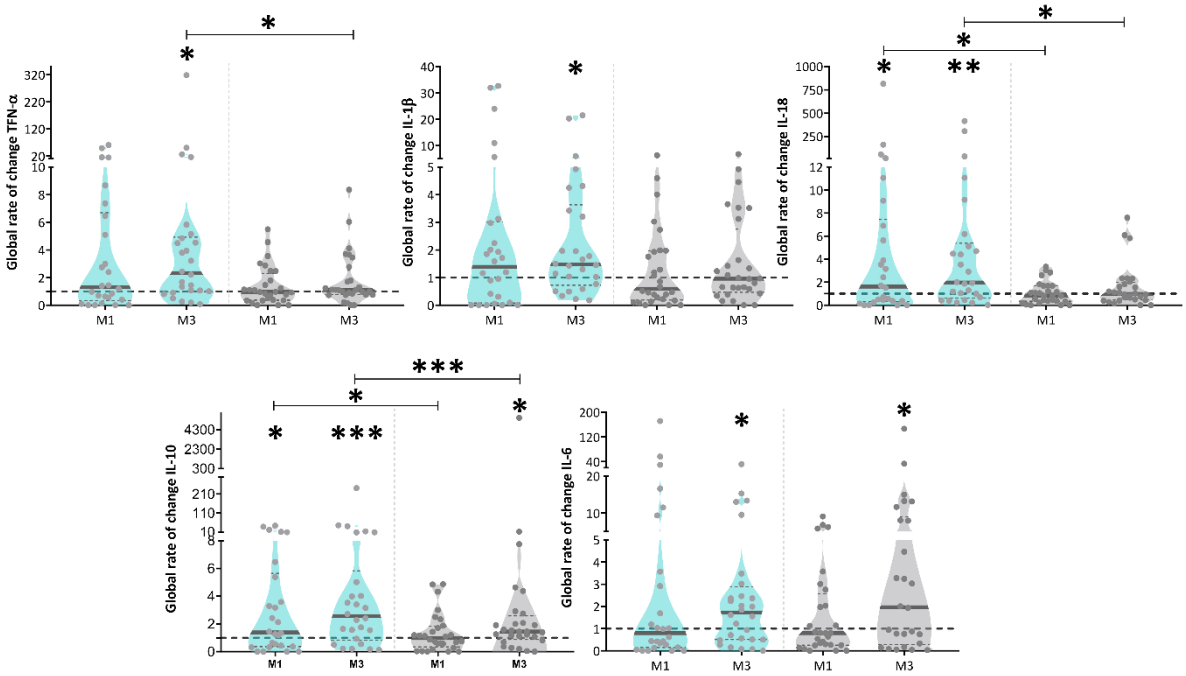
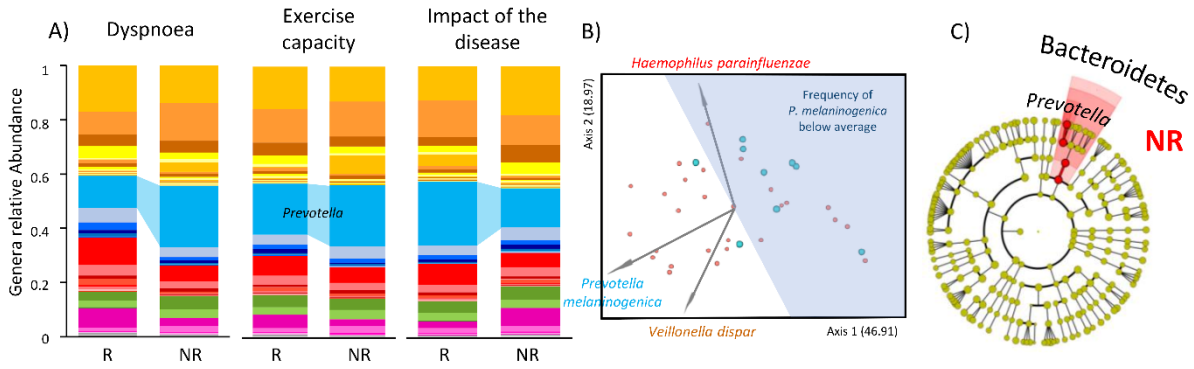


Figure 2. The inflammatory profile of people with COPD changes upon pulmonary rehabilitation. Global rate of change represents the ratio between cytokine values measured at baseline and M1 or M3 (*i.e.* M1/M0 and M3/M0). Differences between M1 and the baseline and between M3 and the baseline were assessed by Wilcoxon signed-rank, in each group (intervention (blue) and control (grey)). Differences between groups at M1 and M3 were assessed by Mann-Whitney U-test. p<0.05, **p<0.01, **p<0.001. All cytokines for which a significant shift was observed are represented in this figure. Supplementary figure 2 shows the data for the remaining 8 cytokines.

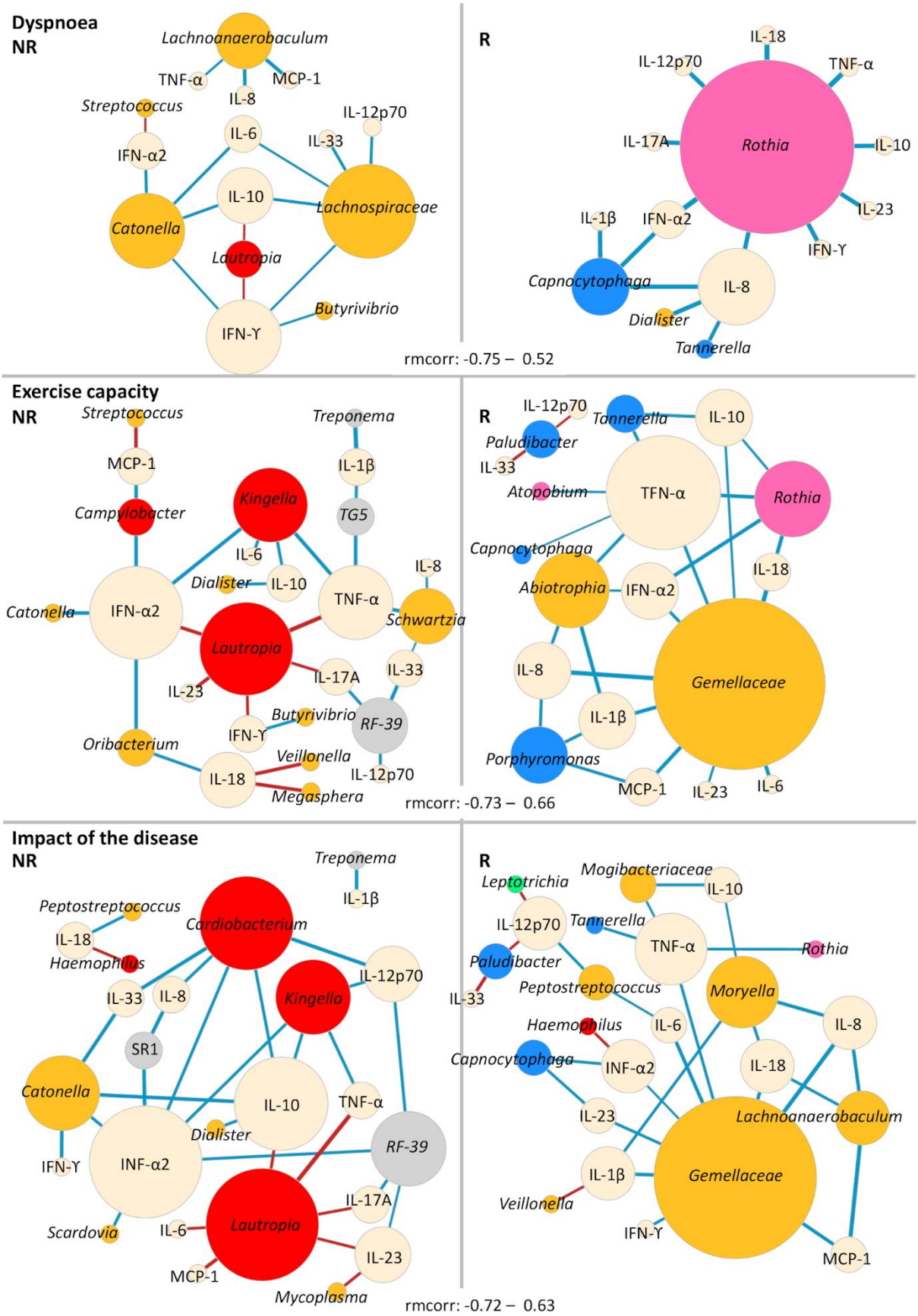
medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.



579

580 **Figure 3. Responders and non-responders present distinct microbiota profiles prior to**
581 **pulmonary rehabilitation. A)** Mean frequency of phyla and genera of bacteria present in
582 Responders (R) and non-responders (NR) to dyspnoea, exercise capacity and impact of the
583 disease. **B)** R and NR to dyspnoea showed distinct microbiota composition prior to PR
584 (PERMANOVA, p=0.04). PCoA analysis using Emperor on Weighted UniFrac distance. The biplot
585 (grey arrows) represent the 3 most relevant ASVs for R and NR segregation. *Prevotella*
586 *melaninogenica* (ASV: d0b698c7298bf04110a6d2f220879bfb) was the major contributor for
587 segregation of R and NR followed by *Veillonella dispar* (ASV:
588 5608c3e6c9de9ceb79610e7786bd0ac) and *Haemophilus parainfluenzae* (ASV:
589 e27680d4009f98f30248d823bc17fb8e). In blue is represented the area where samples
590 presented a mean frequency of *P. melaninogenica* below the average of the dataset. 78% of R
591 showed reduced mean frequencies of *P. melaninogenica* prior to PR while only 37% of NR
592 followed the same trend. **C)** Cladogram highlighting differentially abundant genera between R
593 and NR to dyspnoea prior to PR, inferred by linear discriminant analysis (LEfSe) at a significance
594 cut-off of 3. *Prevotella* is significantly enriched in NR to dyspnoea, with an effect-size of 9.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

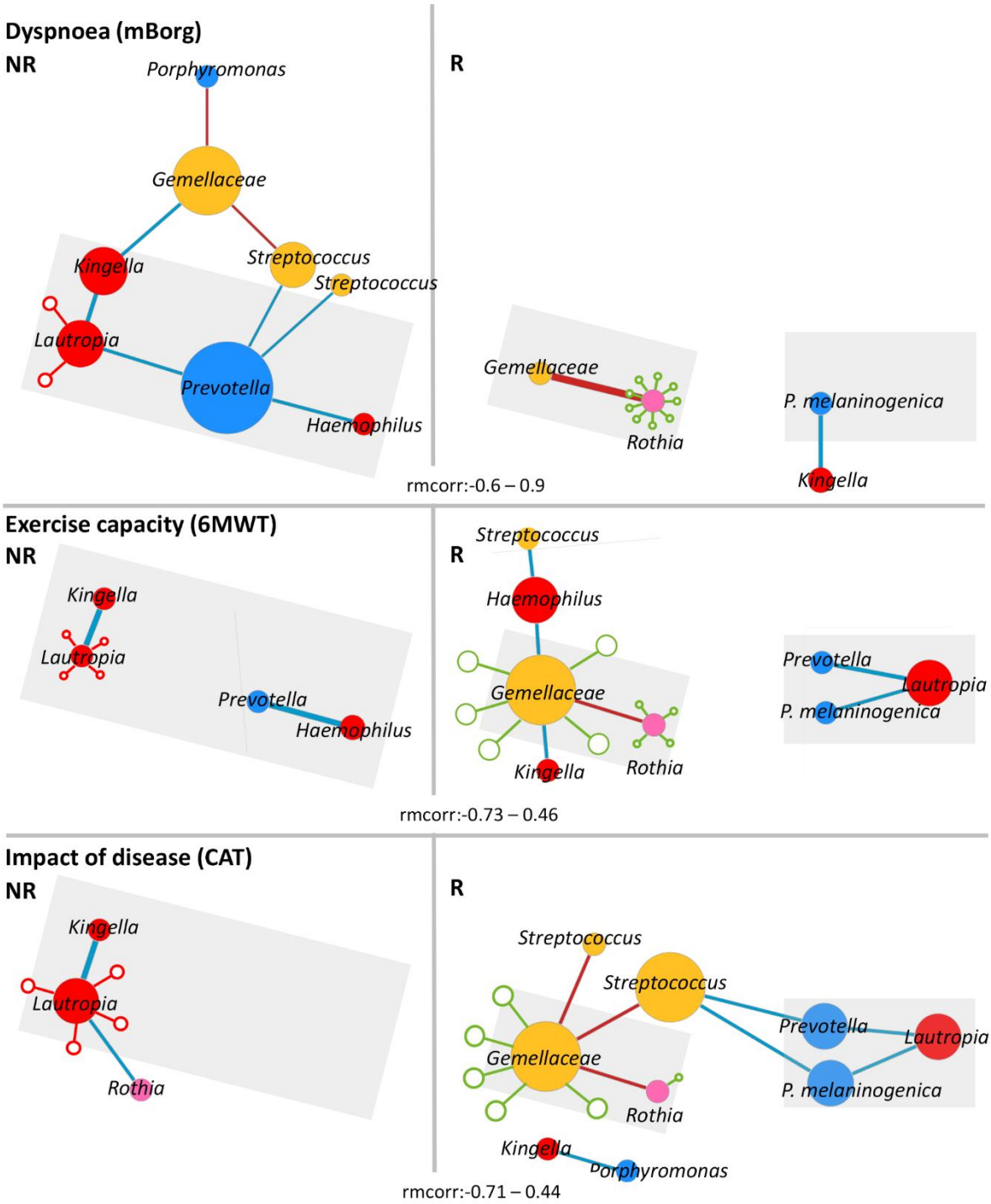


595

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Figure 4. Responders and non-responders to dyspnoea, exercise capacity and impact of the disease present distinct patterns of longitudinal bacteria-inflammatory markers' correlation. Correlation networks representing the repeated-measures correlations between bacteria and inflammatory markers in non-responders and responders to **A)** dyspnoea (rmcorr: [-0.75 – 0.52]) **B)** exercise capacity (rmcorr: [-0.73 – 0.66]) and **C)** impact of the disease (rmcorr: [-0.72 – 0.63]). Prior to the analysis, genera and ASVs: d0b698c7298bf04110a6d2f220879bfb and *a5189f77a2cfeab3bc1602ff5c8ac3e9* relative frequencies were transformed with arcsine square root transformation and inflammatory markers' concentration was transformed with $\log_{10}$. The diameter of the nodes is proportional to the number of connections. Bacterial nodes are coloured with the colour code of each bacterial phyla (Firmicutes: yellow, Bacteroidetes: blue; Proteobacteria: red; Fusobacteria: green Actinobacteria: pink; Other phyla: grey). Inflammatory markers' nodes are represented in beige. Positive correlations are represented in red, while negative correlations are represented in blue. The width of edges is proportional to the correlation coefficient.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.



610

611 **Figure 5. Responders and non-responders to dyspnoea, exercise capacity and impact of the**

612 **disease present distinct patterns of longitudinal bacteria-bacteria correlation.** Correlation

613 networks representing the longitudinal correlations between a subset of bacterial genera and

614 ASVs: d0b698c7298bf04110a6d2f220879bfb and *a5189f77a2cfeab3bc1602ff5c8ac3e9* non-

615 responders and responders to **A)** dyspnoea (rmcorr: [-0.6 – 0.9]), **B)** exercise capacity (rmcorr: [-

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

0.73 – 0.46]), and **C)** impact of the disease (rmcorr: [-0.71 – 0.44]). Grey triangles highlight common patterns of bacterial correlations between domains. Empty blue and red circles represent cytokines with which the correspondent taxon negatively and positively correlates, respectively. These correlations were inferred through the correlation analysis between bacteria-inflammatory markers and are represented here to facilitate the interpretation of the bacteria-bacteria correlation network.

Prior to the analysis, genera and ASVs: d0b698c7298bf04110a6d2f220879bfb and *a5189f77a2cfeab3bc1602ff5c8ac3e9* relative frequencies were transformed with arcsine square root transformation. The diameter of the nodes is proportional to the number of connections. Bacterial nodes were coloured according to the colour code of each bacterial phyla (Firmicutes: yellow, Bacteroidetes: blue; Proteobacteria: red; Fusobacteria: green Actinobacteria: pink; Other phyla: grey). Positive correlations were represented in red, while negative correlations were represented in blue. The width of edges is proportional to the correlation coefficient.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY MATERIAL

Methods

Participants and sample collection

The pulmonary rehabilitation (PR) program was composed by twice a week 60-minute sessions of moderate exercise training and psychoeducational sessions once every other week. Detailed description of the PR program can be found elsewhere [1]. Patients were followed monthly for 5 consecutive months. Thirty-eight patients undertook a 12-week community-based PR program, intervention group, and the remaining 38 integrated the control group. In intervention group the time frame encompasses 1 month before PR, 3 months during PR and 2 months after PR. The PR program was delivered by a multidisciplinary team of healthcare professionals. During this period sociodemographic, anthropometric, clinical data were collected at M0 and M3 (corresponding to pre-post intervention) and saliva samples (monthly, passive drool method) were collected using a structured protocol [1].

Sociodemographic (age, sex, educational level), anthropometric (weight and height to compute body mass index), clinical (smoking habits, number of exacerbations and hospitalizations in the past year, past 3 months and past month, medication used, long-term oxygen, comorbidities - Charlson Comorbidity Index [2], level of airway obstruction-spirometry (FEV_1, FVC, FEV_1pp) (MicroLab 3535, CareFusion, Kent, UK) [3,4], medication including long term oxygen therapy, impact of the disease – COPD Assessment Test (CAT) [5,6], exercise capacity – six-minute walk test (6MWT) [7,8] data were collected following a published structured protocol of the team [1]. Dyspnoea at rest was assessed with the modified Borg Scale (mBorg) which is a 10-item scale. GOLD grades were defined according to FEV1 percentage predicted for each individual. GOLD groups were defined combining the number of exacerbations and hospital admissions of each patient in the year before enrolment with their CAT scores. From the initially predicted 456 samples a total of 418 saliva samples were collected, the number of samples collected per group

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

and timepoint can be found in supplementary table 1. That is, our study has a 9% rate of missing saliva samples, particularly in the last timepoints M4 and M5. This can be largely explained by difficulties in reaching patients and keeping them motivated to collaborate in data collection for a long period.

Response to PR

Response/non-response to PR was determined with the published minimal clinical importance differences (MCIDs) for the modified Borg scale-Dyspnoea (mBorg), 6MWT and COPD assessment test (CAT) were: -1 point [9] , 25m [10] and -2 points [11], respectively.

DNA extraction

Prior to DNA extraction, samples were thawed at room temperature and centrifuged at 10,000xg for 10 minutes. Supernatants were saved for subsequent quantification of inflammatory markers and pellets were used for DNA extraction following QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) protocol with minor modifications: initial sample volume was set to 400µL and the volumes of buffers and Qiagen protease were adjusted. Elution volume was reduced to a quarter of the recommended. Thirty-eight negative controls where saliva was replaced by phosphate-buffered saline were performed in order to control for background bacterial contamination. Quality and quantity of the extracted DNA was assessed in Denovix DS-11 spectrophotometer, with OD260/280 and OD260/230 ratios.

16S rRNA gene amplification and sequencing

16S rRNA gene amplification and sequencing was carried out at the Gene Expression Unit from Instituto Gulbenkian de Ciência following the implemented protocol. Briefly, for each sample, the hypervariable V4 region of 16S rRNA gene was amplified, using universal pair of primers F515 (5'-CACGGTCGKCGGCGCCATT-3') / R806 (5'-GGACTACHVGGGTWTCTAAT-3'). Samples were then pair-end-sequenced on an Illumina MiSeq Benchtop Sequencer, following Illumina recommendations.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Quantification of inflammatory markers in saliva samples

The bead assay LEGENDplex™ Human Inflammation Panel 1 (13-plex) with V-bottom Plate (BioLegend, San Diego, CA, USA), previously used successfully with saliva samples [12,13], was used to quantify inflammatory markers (IL-1β, IFN-α2, IFN-γ, TNF-α, MCP-1, IL-6, IL-8, IL-10, IL-12p70, IL-17A, IL-18, IL-23, and IL-33) in oral samples from intervention and control groups.. Specifically, the supernatants obtained after sample centrifugation for 10 min at 10.000g, from the intervention (at baseline (M0), after 1 month of PR (M1), and after 3 months of PR (M3)) and control groups (in the timepoints M0, M1 and M3), were used for cytokine quantification following manufacturer's recommendations [14]. Data acquisition was performed with BD AccuriTM C6 Plus flow cytometer and analysed with online version of LEGENDplex™ Data Analysis Software Suite [14]. For values below the detection limit an estimate was obtained using the standard curve. For values above the threshold of detection, the value corresponding to the maximum detection limit was used to replace cytokine values.

Microbiota, inflammatory markers, and statistical analyses

Sample characterisation (relative to Table 1)

Descriptive statistics was used to characterize clinical date of intervention and control groups at baseline. Data normality was assessed with Shapiro-Wilk and D'Agostino-Pearson omnibus normality tests, ensuring the assumptions of parametric statistics approach. Comparisons between intervention and control group were conducted with unpaired t-test with Welch's correction, when quantitative data followed normal distribution, Mann-Whitney U-test, when quantitative data violated the assumptions of parametric tests and Chi-square test, when qualitative data was considered (statistical analyses were performed in GraphPad Prism 8 [15] and R software v 3.6.1 [16]). Statistical significance was considered for p-values below 0.05.

Analysis of illumina paired-end reads

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

QIIME2 2020.8 [11,17] was used to perform microbiota analyses. Demultiplexed 16s paired-end sequences were imported and q2-vsearch plugin [18] was applied to join forward and reverse reads. Quality control was assessed via q-score base filtering, chimera removing and 16S-denoising with Deblur [19,20]. Next, to exclude bacterial background contaminations, we used the *DECONTAM* package [21,22] from *R* software [16] with the prevalence method and a threshold of 0.5, meaning that ASVs were classified as contaminants if present in a higher fraction of negative controls than true samples. This allowed us to identify 235 contaminant ASVs (available in supplementary file 2), that together with ASVs from *mitochondria*, *chloroplasts* and *cyanobacteria*were removed from the dataset prior to conducting subsequent analyses. Results from previous steps were summarized in a feature table. Q2-phylogeny plugin [23] was next employed to produce a MAFFT alignment [24] of ASVs which was later used to construct a rooted phylogeny with FastTree2 [25] for subsequent applications.

Taxonomy assignment of ASVs was performed with q2-feature-classifier plugin [26,27], through classify-sklearn method with pre-trained Naïve Bayes classifier against 99%-eHOMD_v15.1 reference database [28] (sequences trimmed to only include 250bp of V4 region, bound by the F515/R806 primer pair).

Categorized samples' metadata used for analyses is supplied in supplementary file 3. Subsequent analyses were performed upon ASVs. Differential abundance analyses were conducted with on both ASVs and OTUs at taxonomic level 6.

Diversity analyses

Alpha- and beta-diversities were estimated with q2-diversity plugin [29] after rarefaction to 8000 sequences per sample (subsample without replacement). Spatial dissimilarities between bacterial communities of different groups and/or periods were assessed with principal coordinate analyses (PCoA) and biplots based on Weighted Unifrac distance. Wilcoxon test and Friedman's test with Dunn's correction were employed to compare alpha diversity among

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

groups, periods and/or timepoints (statistical analyses were performed in *GraphPad Prism 8* [15] and *R stats* package [30] from *R* software [16]). Statistical significance was considered for p-values below 0.05.

Differences in beta-diversity between groups, periods and/or time points were quantified by permutational multivariate analysis of variance (PERMANOVA) [31] conducted with *adonis2* function [32] (*vegan* package [33] from *R* software [16]). For all statistical analyses a p-value of ≤0.05, corrected for multiple testing whenever necessary, was considered statistically significant.

Differential abundance analyses of OTUs

Analysis of composition of microbiomes (ANCOM) [34] and Linear discriminant effect size (LEfSe) analysis [35] were performed to identify differentially abundant OTUs between groups, periods and/or time points. LEfSe is an algorithm for high-dimensional biomarker discovery that uses linear discriminant analysis to estimate the effect size of each taxon and does not account for the compositional nature of the microbiota. ANCOM uses a log ratio analysis to make point estimates of the variance and mean, taking into consideration the compositional nature of the data. These analyses were conducted with the feature table collapsed at genus taxonomic level (L6). LEfSe was performed in the online version [36] with an Linear Discriminant Analysis score of 3 for significance. ANCOM was performed in *R* software with *ANCOM 2.0* script [37] with taxa-wise multiple correction and a W cut-off of significance of 0.7, both recommended by the developer, based on simulation data.

Longitudinal analyses (linear mixed-effects models)

- *Longitudinal differential abundance analyses of OTUs with ANCOM II*

Longitudinal differential abundance analyses were carried out between patients undergoing PR and controls, or R and NR to exercise capacity, dyspnoea and impact of the disease, with analysis of composition of microbiomes (ANCOM II) [38]. Similar to classic ANCOM, ANCOM II uses a log

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

ratio analysis to make point estimates of the variance and mean [39], taking into consideration the compositional nature of the data, but comprises an additional step to deal with different types of zeros, which was further described by Kaul et al [38]. These analyses were conducted with the feature table collapsed at genus taxonomic level (L6) as well as with the feature table not collapsed, e.g., using only ASVs. ANCOM II was performed in R software [16] using ANCOM 2.1 script [40] with taxa-wise multiple correction and a W cut-off of 0.7 significance (both recommended by the developer, based on simulation data).

- *Data normalization and transformation*

To evaluate the global rate of change of beta diversity of patients undergoing PR and controls, all the timepoints of both groups were normalized through subtraction of baseline values.

- Differences in the longitudinal dynamics of OTUs/ASVs' frequencies between groups and between R and NR to exercise capacity, dyspnoea, and impact of the disease were assessed based on the arcsine square root transformed relative frequencies in each timepoint. Additionally, raw data was filtered so that the analysis included only ASVs/OTUs that were present in ≥20% of the samples. *Loess Lines*

Loess Lines plots were produced to observe dissimilar tendencies during the 5 consecutive months (M0 to M5) of the study between patients undergoing PR and controls or R and NR to exercise capacity, dyspnoea and impact of the disease. In brief, Loess regressions were fitted over longitudinal data beta-diversity and ASVs/OTUs' relative frequencies and then plotted with ggplot2 package from R software [16].

- *Linear mixed-effects models*

Longitudinal analyses with linear mixed-effects models, were performed to compare the longitudinal dynamics of beta- diversity (Weighted Unifrac distances) and relative frequencies of genera and 4 ASVs (top 4 ASVs responsible for patients' separation in principal coordinate analysis), between patients undergoing PR and controls. Differences in the longitudinal

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

trajectories of genera/ASVs' relative frequencies between R and NR to exercise capacity, dyspnoea and impact of the disease were also evaluated. Regarding the models for beta diversity, independent variables included the timepoints (5 months) and the experimental design groups (intervention and control). Subjects were incorporated in the model as random factors, using *lmer* function of the *lme4* package [41] of R software [16]. P-values were obtained using the *Anova* function from *R-stats* package [30] and statistically significant results were considered valid if the assumptions of the model were validated.

To assess the effect of PR in the dynamics of each OTU/ASVs, the arcsine square root transformed frequencies of each OTU/ASV were defined as dependent variable in the correspondent model. Moreover, the experimental design groups (intervention and control or R and NR to exercise capacity, dyspnoea and impact of the disease) and timepoints were set as independent variables. Subjects were adjusted as a random factor. P-values were obtained using the *Anova* function and statistically significant results were considered valid if the assumptions of the model were validated. Additionally, contrast analysis with the Bonferroni correction was executed to identify the timepoints in which the mean abundance of relevant OTUs/ASVs is significantly different between R and NR to the several domains. Statistical significance was considered for p-values below 0.05.

Analyses of inflammatory markers

To assess changes of each inflammatory marker, the ratio between each timepoint and the baseline (M1/M0 and M3/M0) was estimated. Wilcoxon Signed-Ranks Test (GraphPad Prism 8) with post-hoc false discovery rate (FDR) correction from multiple comparisons was applied to compare ratios at M1 and M3 to baseline in all experimental groups. Mann-Whitney U-test was used to assess the differences in the median ratio of each inflammatory marker per time-point (M1 vs M1 and M3 vs M3, respectively), among the experimental groups. Z-test for variance (*PairedData* package [42] from *R* software [16]) were executed to assess differences in the

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

variance of ratios of each inflammatory marker per time-point (M1 vs M1 and M3 vs M3, respectively), among the experimental groups. Statistical significance was considered for *p-values* below 0.05.

Repeated Measures Correlations

Repeated measures correlations were performed to assess the longitudinal co-variation of the frequency of genera/ASVs and inflammatory markers during PR, in the groups of R and NR to exercise capacity, dyspnoea and impact of the disease. Correlations were computed with Rmcorr function from *rmcorr* package [43] (R software [16]). Prior to correlation analyses, estimated values of inflammatory markers were transformed with $\log_{10}$ transformation, OTUs list was filtered to consider only OTUs that were present in at least 20% of the samples. Furthermore, the filtered dataset was transformed with arcsine square root transformation. Patients with only two longitudinal measurements, M0 and M3, were also included in the analyses. Statistical significance was considered for p-values below 0.05, however results were only accepted as valid if the assumptions of the *rmcorr* model were validated (namely, model residuals normally distributed and centered in zero). Valid correlations, according to previous criteria, were then plotted in correlation network plots with *igraph* package ([44] from R software [16], with node diameter proportional to the number of correlations and the edges' width proportional to the correlation strength.

The same analysis (including timepoints M0, M1 and M3) with a subset of bacterial genera and ASVs was carried out to assess bacterial interactions that could be related with PR responsiveness. Since compositional data is intrinsically correlated the subgroup of bacterial genera/ASVs included the major hubs found in the correlation between bacteria and inflammatory markers (*Lautropia*, *Rothia*, *Gemellaceae* and *Kingella*) and six other oral taxa previously associated with severity [45] (*Prevotella*, *P. melaninogenica*, *Streptococcus*, *Streptococcus* sp., *Haemophilus* and *Porphyromonas*).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Supplementary Tables

Supplementary Table 1 – Number of saliva samples collected per group in each timepoint over the 5-month period.

Time-points	Intervention (n=38)	Control (n=38)
M0	38	38
M1	37	38
M2	38	38
M3	37	38
M4	22	38
M5	18	38
Total	190	228

Supplementary Table 2 – Variance of the global rate of change estimated for each cytokine at M1 and M3 in the intervention and control groups. Z-test for variance was performed to assess significant differences between the groups.

Time-point	Cytokine	Intervention (n=26)	Control (n=28)	test's statistics	*p-value*
M1	IL-1β	89.5	2.5	36.1	<0.0001
	IFN-α2	149.0	0.2	743.2	<0.0001
	IFN-ϒ	92.4	2.7	34.4	<0.0001
	TNF-α	220.3	2.0	111.4	<0.0001
	MCP-1	1242.3	3.5	357.2	<0.0001
	IL-6	1214.6	5.8	210.3	<0.0001
	IL-8	1796.4	3.5	517.0	<0.0001
	IL-10	126.0	2.0	63.3	<0.0001
	IL-12p70	3.6	2.6	1.4	0.4
	IL-17A	12.5	0.3	45.9	<0.0001
	IL-18	25885.7	0.9	28096.0	<0.0001
	IL-23	180.8	1.4	130.2	<0.0001
	IL-33	3.1	0.5	5.9	<0.0001
M3	IL-1β	29.2	3.0	9.8	<0.0001
	IFN-α2	19.3	0.1	151.9	<0.0001
	IFN-ϒ	61.3	0.4	152.8	<0.0001
	TNF-α	3867.6	3.7	1050.3	<0.0001
	MCP-1	6062.1	737803.0	0.008	<0.0001
	IL-6	828.6	47.7	17.4	<0.0001
	IL-8	1002.9	208.3	4.8	0.0001
	IL-10	2248.1	1085998.9	0.002	<0.0001
	IL-12p70	5.6	0.6	9.5	<0.0001
	IL-17A	28.0	0.3	97.2	<0.0001

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

IL-18	9716.7	3.5	2750.1	<0.0001
IL-23	87.0	652.1	0.1	<0.0001
IL-33	2.2	0.2	13.7	<0.0001

p-values were obtained with Z-test for variance

Supplementary Table 3 – Summary of the effects of pulmonary rehabilitation in people with chronic obstructive pulmonary disease (n=38)

Clinical Parameters	mean Pre (n=38)	SD	mean Post (n=38)	SD	p-value	Mean Difference	Cohen's D
BMI	26.03	4.3	25.8	4.1	0.07	-0.28	0.31
mBorg	0.79	1.2	1.1	1.7	0.5	0.28	0.18
CAT	17.08	8.03	13.5	7.4	0.001	-3.6	0.46
6MWT: walked distance in m	389	132.2	434.4	134.2	0.0006	45.4	0.34

BMI: Body Mass Index; mBorg: modified Borg Scale of Dyspnoea; CAT: COPD assessment test; 6MWT: six-minute walk test

Supplementary Table 4 – Variance of the global rate of change estimated for each cytokine at M1 and M3 in R and NR to dyspnoea (mBorg). Z-test for variance was performed to assess significant differences between R and NR.

Time-point	Cytokine	NR (n=19)	R (n=7)	test's statistics	*p-value*
M1	IL-1β	98.6	76.4	0.8	0.8
	IFN-α2	202.9	1.3	0.006	<0.0001
	IFN-ϒ	123.4	3.0	0.02	0.0002
	TNF-α	290.7	12.4	0.04	0.0008
	MCP-1	10.3	4627.5	447.5	<0.0001
	IL-6	203.6	4113.9	20.2	<0.0001
	IL-8	16.1	6696.4	416.0	<0.0001
	IL-10	163.9	10.0	0.06	0.002
	IL-12p70	4.6	0.2	0.05	0.001
	IL-17A	16.5	0.3	0.02	<0.0001
	IL-18	1436.0	94746.6	66.0	<0.0001
	IL-23	245.6	1.4	0.006	<0.0001
	IL-33	3.8	1.1	0.3	0.14
344M3	IL-1β	20.2	53.2	2.6	0.1
	IFN-α2	24.0	8.1	0.3	0.18
	IFN-ϒ	18.2	161.6	8.9	0.0003
	TNF-α	5239.5	329.0	0.06	0.003
	MCP-1	2.9	22412.2	7796.1	<0.0001
	IL-6	1108.1	128.0	0.1	0.013

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

IL-8	10.1	3717.1	367.9	<0.0001
IL-10	2999.4	369.2	0.1	0.02
IL-12p70	6.6	3.4	0.5	0.42
IL-17A	32.4	18.3	0.6	0.49
IL-18	4862.1	24570.4	5.1	0.007
IL-23	29.7	227.1	7.6	0.0007
IL-33	2.5	1.9	0.8	0.79

p-values were obtained with Z-test for variance

Supplementary Table 5 – Variance of the global rate of change estimated for each cytokine at M1 and M3 in R and NR to exercise capacity (6MWT). Z-test for variance was performed to assess significant differences between R and NR.

Time-point	Cytokine	NR (n=10)	R (n=16)	test's statistics	*p-value*
M1	IL-1β	99.9	88.9	0.9	0.81
	IFN-α2	382.1	2.9	0.008	<0.0001
	IFN-ϒ	199.1	27.8	0.1	0.001
	TNF-α	491.3	19.1	0.04	<0.0001
	MCP-1	17.9	2024.1	113.0	<0.0001
	IL-6	349.4	1811.8	5.2	0.017
	IL-8	29.7	2918.0	98.3	<0.0001
	IL-10	263.4	7.3	0.03	<0.0001
	IL-12p70	7.2	1.1	0.15	0.002
	IL-17A	29.5	1.4	0.05	<0.0001
	IL-18	2534.3	41015.4	16.2	0.0002
	IL-23	467.2	1.9	0.004	<0.0001
	IL-33	5.3	1.7	0.3	0.054
M3	IL-1β	35.3	27.3	0.8	0.63
	IFN-α2	47.7	1.0	0.02	<0.0001
	IFN-ϒ	53.6	68.5	1.3	0.73
	TNF-α	9730.2	15.3	0.002	<0.0001
	MCP-1	5.1	9858.5	1918.8	<0.0001
	IL-6	2081.7	26.6	0.01	<0.0001
	IL-8	17.3	1632.1	94.3	<0.0001
	IL-10	5630.9	100.1	0.02	<0.0001
	IL-12p70	5.5	6.0	1.1	0.93
	IL-17A	65.1	4.7	0.07	<0.0001
	IL-18	9209.6	10663.3	1.2	0.85
	IL-23	189.6	23.9	0.1	0.0005
	IL-33	2.9	2.0	0.7	0.47

p-values were obtained with Z-test for variance

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

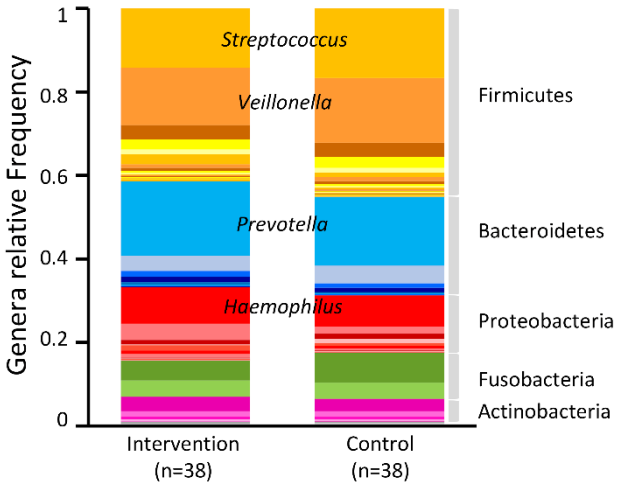
Supplementary Table 6 – Variance of the global rate of change estimated for each cytokine at M1 and M3 in R and NR to the impact of disease (CAT). Z-test for variance was performed to assess significant differences between R and NR.

Time-point	Cytokine	NR (n=12)	R (n=14)	test's statistics	*p-value*
M1	IL-1β	47.3	131.1	2.8	0.099
	IFN-α2	1.0	276.3	287.0	<0.0001
	IFN-ϒ	3.0	168.7	56.2	<0.0001
	TNF-α	19.0	385.9	20.3	<0.0001
	MCP-1	2692.6	14.0	0.005	<0.0001
	IL-6	2401.8	232.5	0.1	0.0002
	IL-8	3883.9	22.1	0.006	<0.0001
	IL-10	40.3	105.9	5.1	0.01
	IL-12p70	0.9	6.1	6.4	0.004
	IL-17A	0.5	22.5	45.2	<0.0001
	IL-18	54805.0	1964.3	0.04	<0.0001
	IL-23	2.1	337.0	164.2	<0.0001
	IL-33	1.9	4.4	2.4	0.16
M3	IL-1β	33.0	27.7	0.8	0.76
	IFN-α2	5.6	32.0	5.7	0.007
	IFN-ϒ	31.5	90.9	2.9	0.09
	TNF-α	227.0	7097.8	31.3	<0.0001
	MCP-1	13133.7	3.5	0.0003	<0.0001
	IL-6	90.2	1473.0	16.3	<0.0001
	IL-8	2174.3	13.2	0.006	<0.0001
	IL-10	160.9	4086.2	25.4	<0.0001
	IL-12p70	5.1	6.4	1.3	0.71
	IL-17A	9.1	45.5	5.0	0.01
	IL-18	14073.7	6639.2	0.5	0.2
	IL-23	124.2	61.6	0.5	0.23
	IL-33	2.8	1.9	0.7	0.48

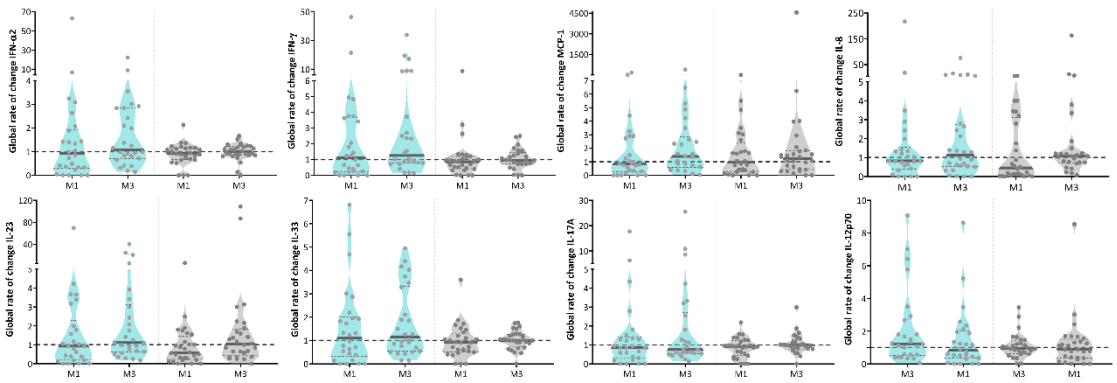
p-values were obtained with Z-test for variance

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Supplementary Figures



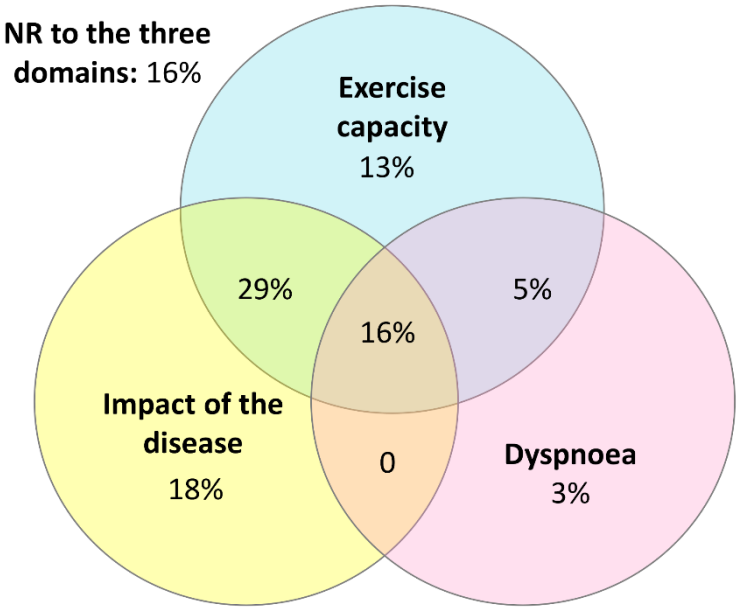
Supplementary Figure 1. Mean frequency of phyla and genera of bacteria present in intervention and control groups at baseline (M0).



Supplementary Figure 2. Violin plots representing the global rate of change of IFN-α2, IFN-ϒ, MCP-1, IL-8, IL-23, IL-33, IL-17A, IL-18 and IL-12p70 in saliva from patients submitted to the 12-week pulmonary rehabilitation programme (blue) and controls (grey). Global rate of change represents the ratio between cytokine values measured at baseline and M1 or M3 (*i.e.* M1/M0 and M3/M0). Differences between M1 and the baseline and between M3 and the baseline were assessed by Wilcoxon signed-rank, in each group (intervention (blue) and control (grey)).

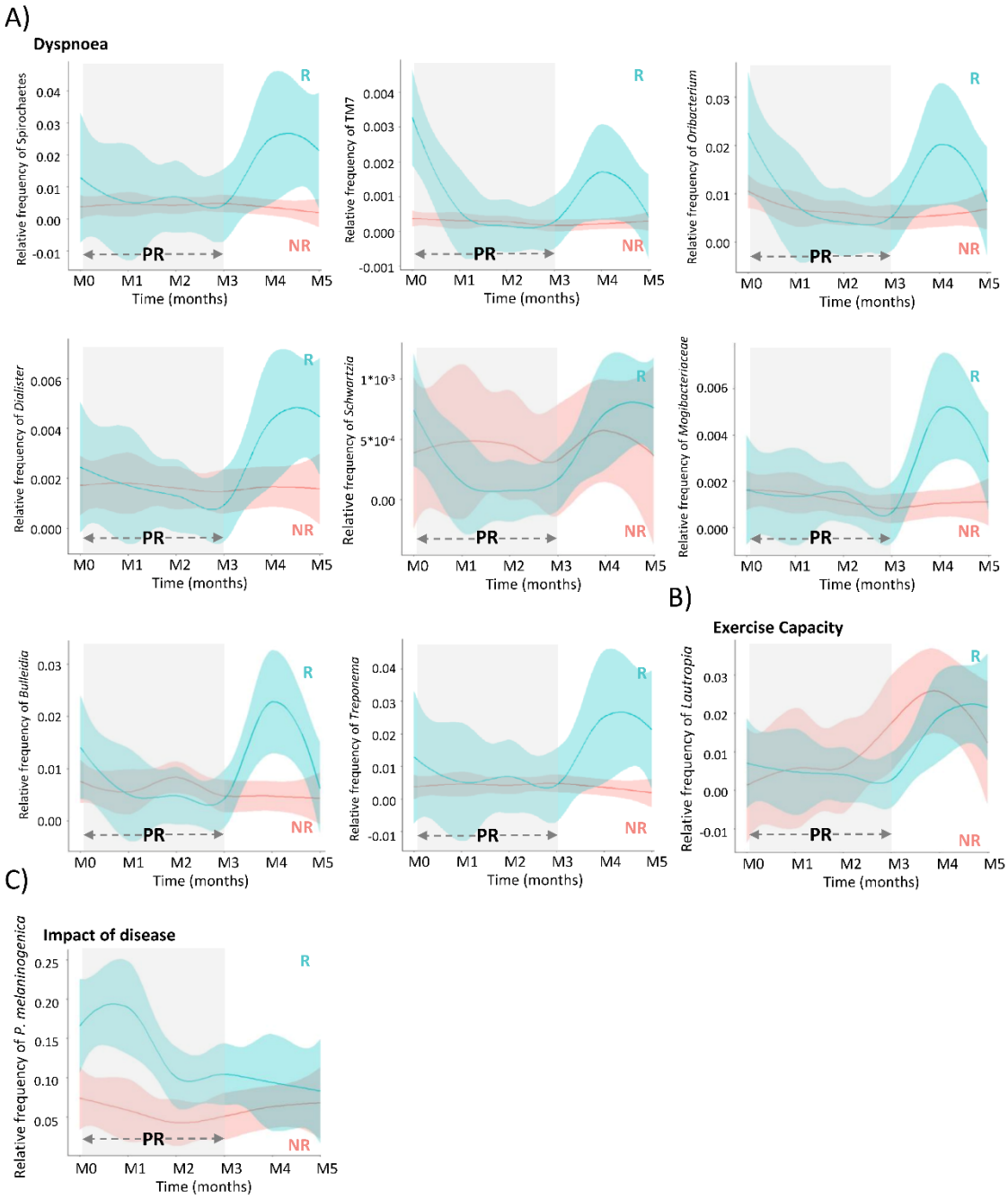
medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Differences between groups at M1 and M3 were assessed by Mann-Whitney U-test. According to these criteria no significant shifts were observed for the cytokines represented in the figure.



Supplementary Figure 3. Venn diagram showing the percentage of overlap between patients' (n=38) responsiveness to each domain: dyspnoea, exercise capacity and impact of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

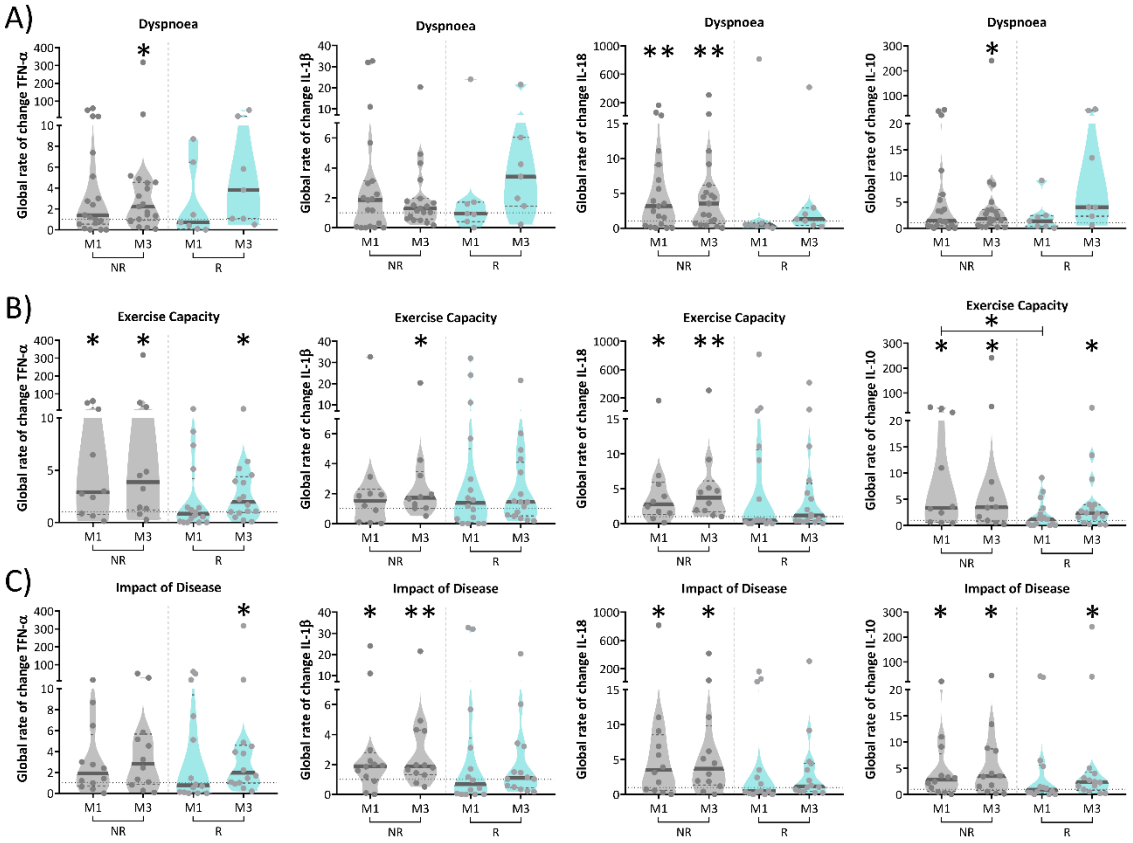


869

870 **Supplementary Figure 4. Relative frequencies over time of taxa presenting significantly**
871 **different dynamics between responders and non-responders.** Linear mixed-effects models
872 were applied to arcsine square root transformed frequencies of ASVs/OTUs to determine the
873 taxa presenting differential dynamics between R and NR to A) dyspnoea during exercise, B)
874 exercise capacity and C) impact of the disease. Grey rectangles include all the time-points where
875 patients were under pulmonary rehabilitation. Blue and red loess lines were fitted over the data
876 points representing the relative frequency of the correspondent taxon in responders and non-

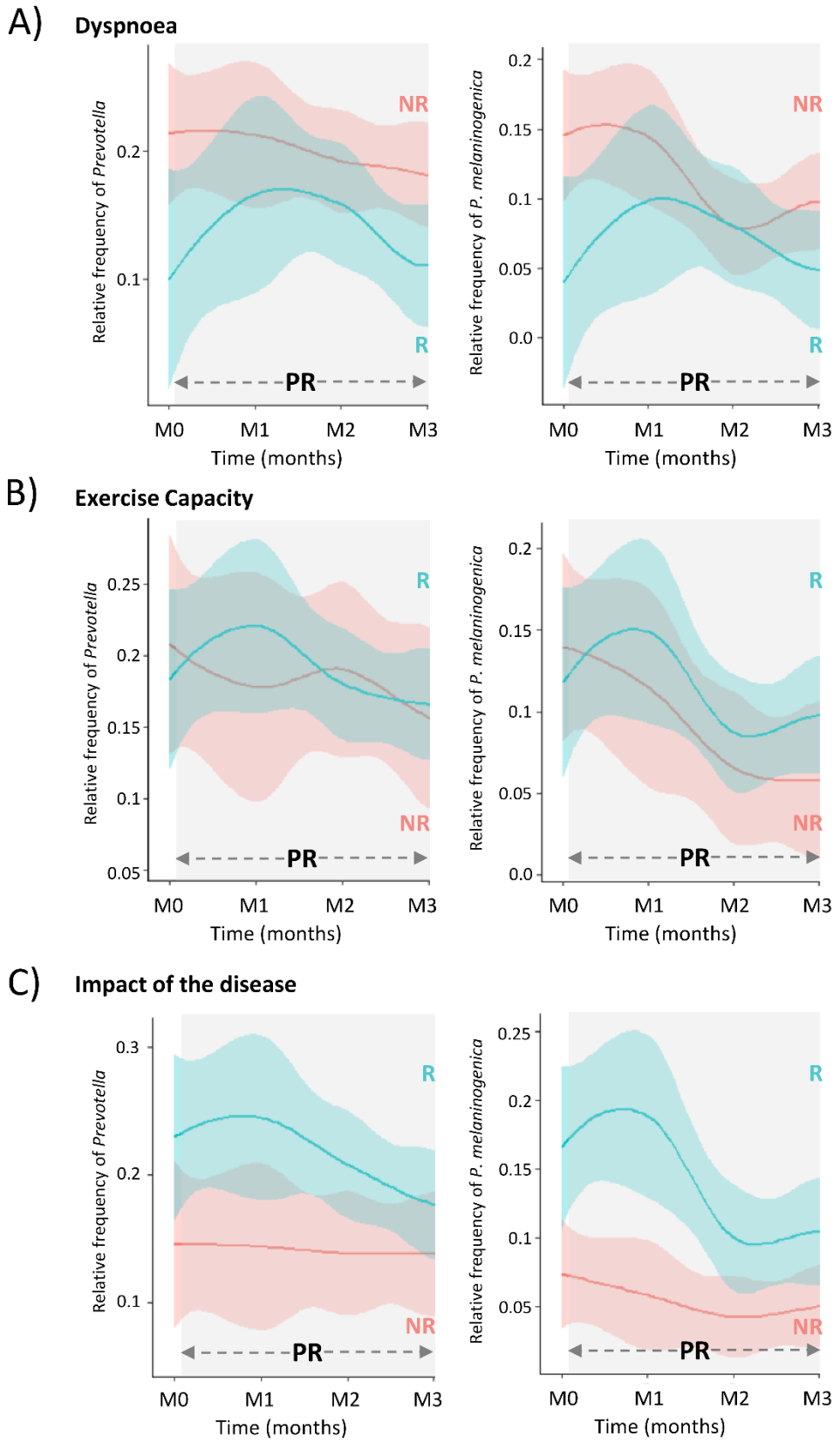
medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

responders (M0=immediately prior to PR; M1, M2, M3= 1,2,3 months after initiating PR; M4, M5=1 and 2 months after terminating PR). The blue and red areas represent the 0.95 confidence intervals.



Supplementary Figure 5. Responsiveness to pulmonary rehabilitation (PR) was associated with specific alterations in the inflammatory markers. Global rate of change represents the ratio between cytokine values measured at baseline and M1 or M3 (*i.e.* M1/M0 and M3/M0). Differences between M1 and the baseline and between M3 and the baseline were assessed by Wilcoxon signed-rank, in each group (intervention (blue) and control (grey)). Differences between groups (R and NR) at M1 and M3 were assessed by Mann-Whitney U-test. *p<0.05, **p<0.01, **p<0.001. A) dyspnoea during exercise, B) exercise capacity and C) impact of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.



889

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Supplementary Figure 6. Frequency of *Prevotella* and *P. melaninogenica* over time in responders and non-responders to A) dyspnoea, B) exercise capacity and C) impact of the disease. The grey rectangles include all the time-points where patients were under pulmonary rehabilitation. Blue and red loess lines were fitted over the data points representing the relative frequency of the correspondent taxon in responders and non-responders (M0=immediately prior to PR; M1, M2, M3= 1,2,3 months after initiating PR). The blue and red areas represent the 0.95 confidence intervals.

Code for data analyses

Qiime2 v20.8

The scripts provided were constructed based on the complete dataset, appropriate selection of samples was performed for analyses considering specific time-points and groups of patients with the command: "*Filter samples*"

Sequencing data import using prepared manifest file:

```
qiime tools import \
 --type 'SampleData[PairedEndSequencesWithQuality]' \
 --input-path pe-33-manifest.tsv \
 --output-path paired-end-demux.qza \
 --input-format PairedEndFastqManifestPhred33
```

Joining of the read pairs, and quality assessment:

```
qiime vsearch join-pairs \
 --p-allowmergestagger \
  --i-demultiplexed-seqs paired-end-demux.qza \
  --o-joined-sequences demux-joined.qza
qiime demux summarize \
  --i-data demux-joined.qza \
  --o-visualization demux-joined.qzv
qiime quality-filter q-score \
  --i-demux demux-joined.qza \
  --o-filtered-sequences demux-joined-filtered.qza \
  --o-filter-stats demux-joined-filter-stats.qza
```

Denoising with Deblur and statistics visualization:

```
qiime deblur denoise-16S \
  --i-demultiplexed-seqs demux-joined-filtered.qza \
  --p-trim-length 250 \
  --p-sample-stats \
  --o-representative-sequences rep-seqs.qza \
  --o-table table.qza \
  --o-stats deblur-stats.qza
qiime feature-table summarize \
  --i-table table.qza \
  --o-visualization table.qzv \
  --m-sample-metadata-file sample-metadata.tsv
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
qiime feature-table tabulate-seqs \
 --i-data rep-seqs.qza \
 --o-visualization rep-seqs.qzv
qiime deblur visualize-stats \
 --i-deblur-stats deblur-stats.qza \
 --o-visualization deblur-stats.qzv
```

Filtering data to remove contaminant ASVs found with DECONTAM:

```
 qiime feature-table filter-features \
--i-table table.qza \
--m-metadata-file features-to-keep.txt \      #features-to-keep.txt is the list of non-contaminant ASVs
--o-filtered-table filt-table.qza
qiime feature-table filter-seqs \
 --i-data rep-seqs.qza \
 --i-table filt-table.qza \
 --o-filtered-data filt-rep-seqs.qza
```

Extract eHOMD_15.1 reference sequences according to primers and sequence length, train Naïve-bayes classifier and preform taxonomy assignment:

```
#Extract reference sequences
qiime tools import \
 --type 'FeatureData[Sequence]' \
 --input-path HOMD_16S_rRNA_RefSeq_V15.1.p9.fasta \
 --output-path HOMD_otus.qza
qiime tools import \
 --type 'FeatureData[Taxonomy]' \
 --input-format HeaderlessTSVTaxonomyFormat \
 --input-path HOMD_16S_rRNA_RefSeq_V15.1.qiime.taxonomy \
 --output-path ref-taxonomy.qza
qiime feature-classifier extract-reads \
 --i-sequences HOMD_otus.qza \
 --p-f-primer GTGCCAGCMGCCGCGGTAA \
 --p-r-primer GGACTACHVGGGTWTCTAAT \
 --p-trunc-len 250 \
 --p-min-length 100 \
 --p-max-length 400 \
 --o-reads ref-seqs.qza
#Train the Naïve-bayes classifier
qiime feature-classifier fit-classifier-naive-bayes \
 --i-reference-reads ref-seqs.qza \
 --i-reference-taxonomy HOMD_taxonomy.qza \
 --o-classifier classifier.qza
#Taxonomy assignment
qiime feature-classifier classify-sklearn \
 --i-classifier HOMD_classifier.qza \
 --i-reads rep-seqs.qza \
 --o-classification taxonomy.qza
qiime feature-classifier classify-sklearn \
 --i-classifier HOMD_classifier.qza \
 --i-reads filt-rep-seqs.qza \
 --o-classification filt-taxonomy.qza
```

Filter samples (e.g. analyses conducted only at M0):

```
 qiime feature-table filter-samples \
--i-table filt-table.qza \
--m-metadata-file samples-to-keep_M0.txt \      #samples-to-keep.txt is the list of samples to retain in each specific analysis
--o-filtered-table M0-table.qza
```

Sequences alignment and generation of a rooted tree for phylogenetic diversity analyses:

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
qiime phylogeny align-to-tree-mafft-fasttree \
 --i-sequences filt-rep-seqs.qza \
 --o-alignment aligned-rep-seqs.qza \
 --o-masked-alignment masked-aligned-rep-seqs.qza \
 --o-tree unrooted-tree.qza \
 --o-rooted-tree rooted-tree.qza
```

Rarefaction and alfa and beta diversity analyses:

```
#Alpha-rarefaction
qiime diversity alpha-rarefaction \
 --i-table filt-table.qza \
 --i-phylogeny rooted-tree.qza \
 --p-max-depth 20000 \
 --m-metadata-file sample-metadata.tsv \
 --o-visualization alpha-rarefaction.qzv
#Diversity analyses
qiime diversity core-metrics-phylogenetic \
 --i-phylogeny rooted-tree.qza \
 --i-table filt-table.qza \
 --p-sampling-depth 8000 \
 --m-metadata-file sample-metadata.tsv \
 --output-dir core-metrics-results
```

Calculation of ASVs relative abundances:

```
qiime feature-table relative-frequency \
--i-table filt-table.qza \
--o-relative-frequency-table freq-table.qza
```

From PCoAs to Biplots:

```
qiime diversity pcoa \
--i-distance-matrix core-metrics-results/weighted_unifrac_distance_matrix.qza \
--p-number-of-dimensions 2 \
--o-pcoa weighted_unifrac_2dpcoa_results.qza
qiime diversity pcoa-biplot \
--i-pcoa weighted_unifrac_2dpcoa_results.qza \
--i-features freq-table.qza \
--o-biplot 2D-weighted_unifrac_biplot_pcoa_results.qza
qiime emperor biplot \
--i-biplot 2D-weighted_unifrac_biplot_pcoa_results.qza \
--m-sample-metadata-file sample-metadata.tsv \
--m-feature-metadata-file filt-taxonomy.qza \
--p-ignore-missing-samples \
--p-number-of-features 3 \
--o-visualization 2D-3-biplot_wU.qzv
```

Taxonomy collapse to genera (L6) level and genera relative abundances calculation:

```
#taxa collapse
qiime taxa collapse \
 --i-table filt-table.qza \
 --i-taxonomy filt-taxonomy.qza \
 --p-level 6 \
 --o-collapsed-table collapsed-filtered-table-l6.qza
#Abundance calculation
qiime feature-table relative-frequency\
 --i-table collapsed-filtered-table-l6.qza \
 --o-relative-frequency-table collapsed-frequency-filtered-table-l6.qza
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Exporting ASVs' table, genera table, genera relative abundance table, alpha-diversity vectors and beta-diversity matrixes:

```
#Exporting the ASVs' table, nr of reads and abundance
qiime tools export \
--input-path filt-table.qza \
--output-path abs-table-ASVs
qiime tools export \
--input-path freq-table.qza \
--output-path rel-table-ASVs
#Exporting the genera table, nr of reads and abundance
qiime tools export \
--input-path collapsed-filtered-table-l6.qza \
--output-path abs-table-L6
qiime tools export \
--input-path collapsed-frequency-filtered-table-l6.qza \
--output-path rel-table-L6
#Converting biom 2.0 table into .txt for lefse
biom convert -i feature-table.biom -o otu_table_rel_l6.tsv --to-tsv --header-key taxonomy
#Exporting alpha diversity vectors
qiime tools extract \
--input-path core-metrics-results/shannon_vector.qza \
--output-path extracted-shannon
qiime tools extract \
--input-path core-metrics-results/faith_pd_vector.qza \
--output-path extracted-faith
#Exporting beta-diversity matrixes
qiime tools extract \
--input-path core-metrics-results/weighted_unifrac_distance_matrix.qza \
--output-path extracted-wu
```

R software

DECONTAM:

```
library(tidyverse)
library(qiime2R)
library(phyloseq)
library(ggplot2)
library(decontam)
?isContaminant
phy<-qza_to_phyloseq(features="E:/PRISMA-final!/PRISMA - SARA/R decontam/table.qza", metadata="E:/PRISMA-final!/PRISMA -
SARA/R decontam/sample-metadata.txt", taxonomy="E:/PRISMA-final!/PRISMA - SARA/R decontam/taxonomy.qza", tmp="C:/tmp
")
head(sample_data(phy))
df <- as.data.frame(sample_data(phy))
df$LibrarySize <- sample_sums(phy)
df <- df[order(df$LibrarySize),]
df$Index <- seq(nrow(df))
ggplot(data=df, aes(x=Index, y=LibrarySize, color=Sample_or_Control)) + geom_point()

##identify contaminants - prevalence
sample_data(phy)$is.neg <- sample_data(phy)$Sample_or_Control == "Control"
contamdf.prev <- isContaminant(phy, method="prevalence", neg="is.neg")
table((contamdf.prev)$contaminant)
head(which(contamdf.prev$contaminant))
head(which(contamdf.prev))
show(contamdf.prev)

contamdf.prev05 <- isContaminant(phy, method="prevalence", neg="is.neg", batch.combine="minimum", threshold=0.5, normaliz
e=TRUE )
table(contamdf.prev05$contaminant)
table(contamdf.prev05)
show(contamdf.prev05)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
# Make phyloseq object of presence-absence in negative controls and true samples
phy.pa <- transform_sample_counts(phy, function(abund) 1*(abund>0))
phy.pa.neg <- prune_samples(sample_data(phy.pa)$Sample_or_Control == "Control", phy.pa)
phy.pa.pos <- prune_samples(sample_data(phy.pa)$Sample_or_Control == "Sample", phy.pa)

# Make data.frame of prevalence in positive and negative samples
df.pa <- data.frame(pa.pos=taxa_sums(phy.pa.pos), pa.neg=taxa_sums(phy.pa.neg),
                    contaminant=contamdf.prev$contaminant)
ggplot(data=df.pa, aes(x=pa.neg, y=pa.pos, color=contaminant)) + geom_point() +
  xlab("Prevalence (Negative Controls)") + ylab("Prevalence (True Samples)")
#writte table
write.table(contamdf.prev05, "E:/PRISMA-final!/PRISMA - SARA/R decontam/contamdf.prev0.5.txt", sep="\t")
```

PCoA and PERMANOVA (example M0 RvsNR to mBorg):

```
library(qiime2R)
library(phyloseq)
library(ggplot2)
library(vegan)
library(readxl)
library(ggordiplots)
library(ggthemes)
library(ggfortify)

wuM0=read.table(file = 'M0distance-matrix.tsv', sep = '\t', header = TRUE)
wuM0_1<- wuM0[,-1]
rownames(wuM0_1)<-wuM0[,1]
phy<-qza_to_phyloseq(features="D:/PRISMA-final!/PcoA spider/table-M0.qza", metadata="D:/PRISMA-final!/PcoA spider/sample-
metadata.tsv", taxonomy="D:/PRISMA-final!/PcoA spider/taxonomy.qza", tree="D:/PRISMA-final!/PcoA spider/rooted-tree.qza", t
mp="C:/tmp")
head(sample_data(phy))
phy
sampledf <- data.frame(sample_data(phy))
PcOA_wU <- cmdscale (wuM0_1, eig = TRUE)
gg_ordiplot(PcOA_wU, groups = sampledf$RBMSD, spiders = TRUE, ellipse = FALSE)
phy_group <- get_variable(phy, "RmBorg")
wu_mBorg<- adonis2(wuM0_1 ~ RmBorg, data = sampledf, method = "wunifrac", by=NULL, permutations = 999)
print(wu_mBorg)
```

ANCOM 2.1 (Intervention vs Control, longitudinal comparison):

```
library(readr)
library(tidyverse)
library(nlme)
library(dplyr)
library(ggplot2)
library(compositions)

# OTU table should be a matrix/data.frame with each feature in rows and sample in columns.
# Metadata should be a matrix/data.frame containing the sample identifier.

# Data Pre-Processing
feature_table_pre_process = function(feature_table, meta_data, sample_var, group_var = NULL,
                                     out_cut = 0.05, zero_cut = 0.90, lib_cut, neg_lb){
  feature_table = data.frame(feature_table, check.names = FALSE)
  meta_data = data.frame(meta_data, check.names = FALSE)
  # Drop unused levels
  meta_data[] = lapply(meta_data, function(x) if(is.factor(x)) factor(x) else x)
  # Match sample IDs between metadata and feature table
  sample_ID = intersect(meta_data[, sample_var], colnames(feature_table))
  feature_table = feature_table[, sample_ID]
  meta_data = meta_data[match(sample_ID, meta_data[, sample_var]), ]

  # 1. Identify outliers within each taxon
  if (!is.null(group_var)) {
    group = meta_data[, group_var]
    z = feature_table + 1 # Add pseudo-count (1)
    f = log(z); f[f == 0] = NA; f = colMeans(f, na.rm = T)
    f_fit = lm(f ~ group)
    e = rep(0, length(f)); e[!is.na(group)] = residuals(f_fit)
    y = t(t(z) - e)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
  outlier_check = function(x){
    # Fitting the mixture model using the algorithm of Peddada, S. Das, and JT Gene Hwang (2002)
    mu1 = quantile(x, 0.25, na.rm = T); mu2 = quantile(x, 0.75, na.rm = T)
    sigma1 = quantile(x, 0.75, na.rm = T) - quantile(x, 0.25, na.rm = T); sigma2 = sigma1
    pi = 0.75
    n = length(x)
    epsilon = 100
    tol = 1e-5
    score = pi*dnorm(x, mean = mu1, sd = sigma1)/((1 - pi)*dnorm(x, mean = mu2, sd = sigma2))
    while (epsilon > tol) {
      grp1_ind = (score >= 1)
      mu1_new = mean(x[grp1_ind]); mu2_new = mean(x[!grp1_ind])
      sigma1_new = sd(x[grp1_ind]); if(is.na(sigma1_new)) sigma1_new = 0
      sigma2_new = sd(x[!grp1_ind]); if(is.na(sigma2_new)) sigma2_new = 0
      pi_new = sum(grp1_ind)/n

      para = c(mu1_new, mu2_new, sigma1_new, sigma2_new, pi_new)
      if(any(is.na(para))) break

      score = pi_new * dnorm(x, mean = mu1_new, sd = sigma1_new)/
        ((1-pi_new) * dnorm(x, mean = mu2_new, sd = sigma2_new))

      epsilon = sqrt((mu1 - mu1_new)^2 + (mu2 - mu2_new)^2 +
                  (sigma1 - sigma1_new)^2 + (sigma2 - sigma2_new)^2 + (pi - pi_new)^2)
      mu1 = mu1_new; mu2 = mu2_new; sigma1 = sigma1_new; sigma2 = sigma2_new; pi = pi_new
    }

    if(mu1 + 1.96 * sigma1 < mu2 - 1.96 * sigma2){
      if(pi < out_cut){
        out_ind = grp1_ind
      }else if(pi > 1 - out_cut){
        out_ind = (!grp1_ind)
      }else{
        out_ind = rep(FALSE, n)
      }
    }else{
      out_ind = rep(FALSE, n)
    }
    return(out_ind)
  }
  out_ind = matrix(FALSE, nrow = nrow(feature_table), ncol = ncol(feature_table))
  out_ind[, !is.na(group)] = t(apply(y, 1, function(i)
    unlist(tapply(i, group, function(j) outlier_check(j)))))

  feature_table[out_ind] = NA
}

# 2. Discard taxa with zeros  >=  zero_cut
zero_prop = apply(feature_table, 1, function(x) sum(x == 0, na.rm = T)/length(x[!is.na(x)]))
taxa_del = which(zero_prop >= zero_cut)
if(length(taxa_del) > 0){
  feature_table = feature_table[- taxa_del, ]
}

# 3. Discard samples with library size < lib_cut
lib_size = colSums(feature_table, na.rm = T)
if(any(lib_size < lib_cut)){
  subj_del = which(lib_size < lib_cut)
  feature_table = feature_table[, - subj_del]
  meta_data = meta_data[- subj_del, ]
}

# 4. Identify taxa with structure zeros
if (!is.null(group_var)) {
  group = factor(meta_data[, group_var])
  present_table = as.matrix(feature_table)
  present_table[is.na(present_table)] = 0
  present_table[present_table != 0] = 1

  p_hat = t(apply(present_table, 1, function(x)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
    unlist(tapply(x, group, function(y) mean(y, na.rm = T)))))
  samp_size = t(apply(feature_table, 1, function(x)
    unlist(tapply(x, group, function(y) length(y[!is.na(y)])))))
  p_hat_lo = p_hat - 1.96 * sqrt(p_hat * (1 - p_hat)/samp_size)

  struc_zero = (p_hat == 0) * 1
  # Whether we need to classify a taxon into structural zero by its negative lower bound?
  if(neg_lb) struc_zero[p_hat_lo <= 0] = 1

  # Entries considered to be structural zeros are set to be 0s
  struc_ind = struc_zero[, group]
  feature_table = feature_table * (1 - struc_ind)

  colnames(struc_zero) = paste0("structural_zero (", colnames(struc_zero), ")")
 }else{
  struc_zero = NULL
 }

 # 5. Return results
 res = list(feature_table = feature_table, meta_data = meta_data, structure_zeros = struc_zero)
 return(res)
}

# ANCOM main function
ANCOM = function(feature_table, meta_data, struc_zero = NULL, main_var, p_adj_method = "BH",
           alpha = 0.05, adj_formula = NULL, rand_formula = NULL, ...){
 # OTU table transformation:
 # (1) Discard taxa with structural zeros (if any); (2) Add pseudocount (1) and take logarithm.
 if (!is.null(struc_zero)) {
  num_struc_zero = apply(struc_zero, 1, sum)
  comp_table = feature_table[num_struc_zero == 0, ]
 }else{
  comp_table = feature_table
 }
 comp_table = log(as.matrix(comp_table) + 1)
 n_taxa = dim(comp_table)[1]
 taxa_id = rownames(comp_table)
 n_samp = dim(comp_table)[2]

 # Determine the type of statistical test and its formula.
 if (is.null(rand_formula) & is.null(adj_formula)) {
  # Basic model
  # Whether the main variable of interest has two levels or more?
  if (length(unique(meta_data%>%pull(main_var))) == 2) {
   # Two levels: Wilcoxon rank-sum test
   tfun = stats::wilcox.test
  } else{
   # More than two levels: Kruskal-Wallis test
   tfun = stats::kruskal.test
  }
  # Formula
  tformula = formula(paste("x ~", main_var, sep = " "))
 }else if (is.null(rand_formula) & !is.null(adj_formula)) {
  # Model: ANOVA
  tfun = stats::aov
  # Formula
  tformula = formula(paste("x ~", main_var, "+", adj_formula, sep = " "))
 }else if (!is.null(rand_formula)) {
  # Model: Mixed-effects model
  tfun = nlme::lme
  # Formula
  if (is.null(adj_formula)) {
   # Random intercept model
   tformula = formula(paste("x ~", main_var))
  }else {
   # Random coefficients/slope model
   tformula = formula(paste("x ~", main_var, "+", adj_formula))
  }
 }

 # Calculate the p-value for each pairwise comparison of taxa.
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
p_data = matrix(NA, nrow = n_taxa, ncol = n_taxa)
colnames(p_data) = taxa_id
rownames(p_data) = taxa_id
for (i in 1:(n_taxa - 1)) {
  # Loop through each taxon.
  # For each taxon i, additive log ratio (alr) transform the OTU table using taxon i as the reference.
  # e.g. the first alr matrix will be the log abundance data (comp_table) recursively subtracted
  # by the log abundance of 1st taxon (1st column) column-wisely, and remove the first i columns since:
  # the first (i - 1) columns were calculated by previous iterations, and
  # the i^th column contains all zeros.
  alr_data = apply(comp_table, 1, function(x) x - comp_table[i, ])
  # apply(...) allows crossing the data in a number of ways and avoid explicit use of loop constructs.
  # Here, we basically want to iteratively subtract each column of the comp_table by its i^th column.
  alr_data = alr_data[, - (1:i), drop = FALSE]
  n_lr = dim(alr_data)[2] # number of log-ratios (lr)
  alr_data = cbind(alr_data, meta_data) # merge with the metadata

  # P-values
  if (is.null(rand_formula) & is.null(adj_formula)) {
    p_data[-(1:i), i] = apply(alr_data[, 1:n_lr, drop = FALSE], 2, function(x){
      suppressWarnings(tfun(tformula,
                    data = data.frame(x, alr_data,
                                      check.names = FALSE))$p.value)
    }
    )
  }else if (is.null(rand_formula) & !is.null(adj_formula)) {
    p_data[-(1:i), i] = apply(alr_data[, 1:n_lr, drop = FALSE], 2, function(x){
      fit = tfun(tformula,
             data = data.frame(x, alr_data, check.names = FALSE),
             na.action = na.omit)
      summary(fit)[[1]][main_var, "Pr(>F)"]
    }
    )
  }else if (!is.null(rand_formula)) {
    p_data[-(1:i), i] = apply(alr_data[, 1:n_lr, drop = FALSE], 2, function(x){
      fit = tfun(fixed = tformula,
             data = data.frame(x, alr_data, check.names = FALSE),
             random = formula(rand_formula),
             na.action = na.omit, ...)
      anova(fit)[main_var, "p-value"]
    }
    )
  }
}
# Complete the p-value matrix.
# What we got from above iterations is a lower triangle matrix of p-values.
p_data[upper.tri(p_data)] = t(p_data)[upper.tri(p_data)]
diag(p_data) = 1 # let p-values on diagonal equal to 1
p_data[is.na(p_data)] = 1 # let p-values of NA equal to 1

# Multiple comparisons correction.
q_data = apply(p_data, 2, function(x) p.adjust(x, method = p_adj_method))

# Calculate the W statistic of ANCOM.
# For each taxon, count the number of q-values < alpha.
W = apply(q_data, 2, function(x) sum(x < alpha))

# Organize outputs
out_comp = data.frame(taxa_id, W, row.names = NULL, check.names = FALSE)
# Declare a taxon to be differentially abundant based on the quantile of W statistic.
# We perform (n_taxa - 1) hypothesis testings on each taxon, so the maximum number of rejections is (n_taxa - 1).
out_comp = out_comp%>%mutate(detected_0.9 = ifelse(W > 0.9 * (n_taxa -1), TRUE, FALSE),
                             detected_0.8 = ifelse(W > 0.8 * (n_taxa -1), TRUE, FALSE),
                             detected_0.7 = ifelse(W > 0.7 * (n_taxa -1), TRUE, FALSE),
                             detected_0.6 = ifelse(W > 0.6 * (n_taxa -1), TRUE, FALSE))

# Taxa with structural zeros are automatically declared to be differentially abundant
if (!is.null(struc_zero)){
  out = data.frame(taxa_id = rownames(struc_zero), W = Inf, detected_0.9 = TRUE,
                   detected_0.8 = TRUE, detected_0.7 = TRUE, detected_0.6 = TRUE,
                   row.names = NULL, check.names = FALSE)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
  out[match(taxa_id, out$taxa_id), ] = out_comp
 }else{
  out = out_comp
 }

 # Draw volcano plot
 # Calculate clr
 clr_table = apply(feature_table, 2, clr)
 # Calculate clr mean difference
 eff_size = apply(clr_table, 1, function(y)
  lm(y ~ x, data = data.frame(y = y,
                              x = meta_data %>% pull(main_var),
                              check.names = FALSE))$coef[-1])

 if (is.matrix(eff_size)){
  # Data frame for the figure
  dat_fig = data.frame(taxa_id = out$taxa_id, t(eff_size), y = out$W, check.names = FALSE) %>%
   mutate(zero_ind = factor(ifelse(is.infinite(y), "Yes", "No"), levels = c("Yes", "No"))) %>%
   gather(key = group, value = x, rownames(eff_size))
  # Replcace "x" to the name of covariate
  dat_fig$group = sapply(dat_fig$group, function(x) gsub("x", paste0(main_var, " = "), x))
  # Replace Inf by (n_taxa - 1) for structural zeros
  dat_fig$y = replace(dat_fig$y, is.infinite(dat_fig$y), n_taxa - 1)

  fig = ggplot(data = dat_fig) + aes(x = x, y = y) +
   geom_point(aes(color = zero_ind)) +
   facet_wrap(~ group) +
   labs(x = "CLR mean difference", y = "W statistic") +
   scale_color_discrete(name = "Structural zero", drop = FALSE) +
   theme_bw() +
   theme(plot.title = element_text(hjust = 0.5), legend.position = "top",
         strip.background = element_rect(fill = "white"))
  fig
 } else{
  # Data frame for the figure
  dat_fig = data.frame(taxa_id = out$taxa_id, x = eff_size, y = out$W) %>%
   mutate(zero_ind = factor(ifelse(is.infinite(y), "Yes", "No"), levels = c("Yes", "No")))
  # Replace Inf by (n_taxa - 1) for structural zeros
  dat_fig$y = replace(dat_fig$y, is.infinite(dat_fig$y), n_taxa - 1)

  fig = ggplot(data = dat_fig) + aes(x = x, y = y) +
   geom_point(aes(color = zero_ind)) +
   labs(x = "CLR mean difference", y = "W statistic") +
   scale_color_discrete(name = "Structural zero", drop = FALSE) +
   theme_bw() +
   theme(plot.title = element_text(hjust = 0.5), legend.position = "top")
  fig
 }

 res = list(out = out, fig = fig)
 return(res)
}

###INTERVENTION vs CONTROL
#importing data#
otu_data = read_tsv("D:/PRISMA-final!/ANCOM 2.1/INTvsCNT.tsv", skip = 1)
otu_id = otu_data$`feature-id`
otu_data = data.frame(otu_data[, -1], check.names = FALSE)
rownames(otu_data) = otu_id

meta_data = read_tsv("D:/PRISMA-final!/ANCOM 2.1/sample-metadata.tsv")[-1, ]
meta_data = meta_data %>% rename(Sample.ID = `#SampleID`)

#data processing#
feature_table = otu_data; sample_var = "Sample.ID"; group_var = "Group"
out_cut = 0.05; zero_cut = 0.90; lib_cut = 8000; neg_lb = TRUE
prepro = feature_table_pre_process(feature_table, meta_data, sample_var, group_var,
                                   out_cut, zero_cut, lib_cut, neg_lb)
feature_table = prepro$feature_table # Preprocessed feature table
meta_data = prepro$meta_data # Preprocessed metadata
struc_zero = prepro$structure_zeros # Structural zero info
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
# Step 2: ANCOM

main_var = "Group"; p_adj_method = "fdr"; alpha = 0.05
adj_formula = NULL; rand_formula = "~ 1 | Subject"
control = lmeControl(maxIter = 100, msMaxIter = 100, opt = "optim")
t_start = Sys.time()
res = ANCOM(feature_table, meta_data, struc_zero, main_var, p_adj_method,
        alpha, adj_formula, rand_formula, control = control)
t_end = Sys.time()
t_run = t_end - t_start # around 30s

write_csv(res$out, "D:/PRISMA-final!/ANCOM 2.1/results_INT_CTR.csv")

# Step 3: Volcano Plot

# Number of taxa except structural zeros
n_taxa = ifelse(is.null(struc_zero), nrow(feature_table), sum(apply(struc_zero, 1, sum) == 0))
# Cutoff values for declaring differentially abundant taxa
cut_off = c(0.9 * (n_taxa -1), 0.8 * (n_taxa -1), 0.7 * (n_taxa -1), 0.6 * (n_taxa -1))
names(cut_off) = c("detected_0.9", "detected_0.8", "detected_0.7", "detected_0.6")

# Annotation data
dat_ann = data.frame(x = min(res$fig$data$x), y = cut_off["detected_0.7"], label = "W[0.7]")

fig = res$fig +
  geom_hline(yintercept = cut_off["detected_0.7"], linetype = "dashed") +
  geom_text(data = dat_ann, aes(x = x, y = y, label = label),
        size = 4, vjust = -0.5, hjust = 0, color = "orange", parse = TRUE)
fig
```

Formatting beta-diversity matrix, global rate of change and Linear mixed effect model between Intervention and Control groups

```
 library(readr)
library(readr)
beta <- read_delim("weighted-unifrac-filtered-distance-matrix.tsv",
            "\t", escape_double = FALSE, trim_ws = TRUE)
View(beta)

beta=beta[,-1]
rownames(beta)=colnames(beta)

##########################################
##### Matrizes com os dados todos #######
##########################################

########
## M0 ##
########

library(tidyverse)
M0=matches(c("M0"), vars=colnames(beta1))

betaM0=beta1[,M0]
rownames(betaM0)=rownames(beta1)
View(betaM0)

M0R=matches(c(
  "sample3R451M0",
  "sample3R452M0",
  "sample3R455M0",
  "sample508M0",
  "sample509M0",
  "sample519M0",
  "sample526M0",
  "sample527M0",
  "sample540M0",
  "sample578M0",
  "sample583M0",
  "sample649M0",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
 "sample663M0",
 "sample665M0",
 "sample676M0",
 "sample678M0",
 "sample694M0",
 "sample695M0",
 "sample794M0",
 "sample795M0",
 "sample801M0",
 "sample807M0",
 "sample809M0",
 "sample830M0",
 "sample833M0",
 "sample837M0",
 "sample851M0",
 "sample880M0",
 "sample884M0",
 "sample885M0",
 "sample891M0",
 "sample897M0",
 "sample907M0",
 "sample934M0",
 "sample935M0",
 "sampleP101M0",
 "sampleP201M0",
 "sampleP204M0"), vars=colnames(betaM0))

betaM0=betaM0[,-M0R]

M0C=matches(c("sample3R451M0",
        "sample3R452M0",
        "sample3R455M0",
        "sample508M0",
        "sample509M0",
        "sample519M0",
        "sample526M0",
        "sample527M0",
        "sample540M0",
        "sample578M0",
        "sample583M0",
        "sample649M0",
        "sample663M0",
        "sample665M0",
        "sample676M0",
        "sample678M0",
        "sample694M0",
        "sample695M0",
        "sample794M0",
        "sample795M0",
        "sample801M0",
        "sample807M0",
        "sample809M0",
        "sample830M0",
        "sample833M0",
        "sample837M0",
        "sample851M0",
        "sample880M0",
        "sample884M0",
        "sample885M0",
        "sample891M0",
        "sample897M0",
        "sample907M0",
        "sample934M0",
        "sample935M0",
        "sampleP101M0",
        "sampleP201M0",
        "sampleP204M0"), vars=colnames(beta1))

betaM0=cbind(betaM0,beta1[,M0C])
betaM0=as.data.frame(t(betaM0))
colnames(betaM0)=colnames(beta1)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
#Dados originais#
loess.linesM0=NULL

for (i in 1:nrow(betaM0)) {
  loess.linesM0=c(loess.linesM0,betaM0[i,])
}

loess.linesM0=as.data.frame(loess.linesM0)
loess.linesM0=t(loess.linesM0)

TPM0=c(rep(3,length(loess.linesM0)))

DEM0=c(rep("I",34*381),rep("C",37*381))


########
## M1 ##
########

M1=matches(c("M1"), vars=colnames(beta1))

betaM1=beta1[,M1]
rownames(betaM1)=rownames(beta1)
View(betaM1)

M1R=matches(c("sample3R451M1",
          "sample3R452M1",
          "sample3R455M1",
          "sample508M1",
          "sample509M1",
          "sample519M1",
          "sample526M1",
          "sample527M1",
          "sample540M1",
          "sample578M1",
          "sample583M1",
          "sample649M1",
          "sample663M1",
          "sample665M1",
          "sample676M1",
          "sample678M1",
          "sample694M1",
          "sample695M1",
          "sample794M1",
          "sample795M1",
          "sample801M1",
          "sample807M1",
          "sample809M1",
          "sample830M1",
          "sample833M1",
          "sample837M1",
          "sample851M1",
          "sample880M1",
          "sample884M1",
          "sample885M1",
          "sample891M1",
          "sample897M1",
          "sample907M1",
          "sample934M1",
          "sample935M1",
          "sampleP101M1",
          "sampleP201M1",
          "sampleP204M1"), vars=colnames(betaM1))

betaM1=betaM1[,-M1R]

M1C=matches(c("sample3R451M1",
          "sample3R452M1",
          "sample3R455M1",
          "sample508M1",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
        "sample509M1",
        "sample519M1",
        "sample526M1",
        "sample527M1",
        "sample540M1",
        "sample578M1",
        "sample583M1",
        "sample649M1",
        "sample663M1",
        "sample665M1",
        "sample676M1",
        "sample678M1",
        "sample694M1",
        "sample695M1",
        "sample794M1",
        "sample795M1",
        "sample801M1",
        "sample807M1",
        "sample809M1",
        "sample830M1",
        "sample833M1",
        "sample837M1",
        "sample851M1",
        "sample880M1",
        "sample884M1",
        "sample885M1",
        "sample891M1",
        "sample897M1",
        "sample907M1",
        "sample934M1",
        "sample935M1",
        "sampleP101M1",
        "sampleP201M1",
        "sampleP204M1"), vars=colnames(beta1))

betaM1=cbind(betaM1,beta1[,M1C])
betaM1=as.data.frame(t(betaM1))
colnames(betaM1)=colnames(beta1)

#Dados originais#
loess.linesM1=NULL

for (i in 1:nrow(betaM1)) {
  loess.linesM1=c(loess.linesM1,betaM1[i,])
}

loess.linesM1=as.data.frame(loess.linesM1)
loess.linesM1=t(loess.linesM1)

TPM1=c(rep(4,length(loess.linesM1)))

DEM1=c(rep("I",34*381),rep("C",34*381))


########
## M2 ##
########

M2=matches(c("M2"), vars=colnames(beta1))

betaM2=beta1[,M2]
rownames(betaM2)=rownames(beta1)
View(betaM2)

M2R=matches(c("sample3R451M2",
        "sample3R452M2",
        "sample3R455M2",
        "sample508M2",
        "sample509M2",
        "sample519M2",
        "sample526M2",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
        "sample527M2",
        "sample540M2",
        "sample578M2",
        "sample583M2",
        "sample649M2",
        "sample663M2",
        "sample665M2",
        "sample676M2",
        "sample678M2",
        "sample694M2",
        "sample695M2",
        "sample794M2",
        "sample795M2",
        "sample801M2",
        "sample807M2",
        "sample809M2",
        "sample830M2",
        "sample833M2",
        "sample837M2",
        "sample851M2",
        "sample880M2",
        "sample884M2",
        "sample885M2",
        "sample891M2",
        "sample897M2",
        "sample907M2",
        "sample934M2",
        "sample935M2",
        "sampleP101M2",
        "sampleP201M2",
        "sampleP204M2"), vars=colnames(betaM2))

betaM2=betaM2[,-M2R]

M2C=matches(c("sample3R451M2",
        "sample3R452M2",
        "sample3R455M2",
        "sample508M2",
        "sample509M2",
        "sample519M2",
        "sample526M2",
        "sample527M2",
        "sample540M2",
        "sample578M2",
        "sample583M2",
        "sample649M2",
        "sample663M2",
        "sample665M2",
        "sample676M2",
        "sample678M2",
        "sample694M2",
        "sample695M2",
        "sample794M2",
        "sample795M2",
        "sample801M2",
        "sample807M2",
        "sample809M2",
        "sample830M2",
        "sample833M2",
        "sample837M2",
        "sample851M2",
        "sample880M2",
        "sample884M2",
        "sample885M2",
        "sample891M2",
        "sample897M2",
        "sample907M2",
        "sample934M2",
        "sample935M2",
        "sampleP101M2",
        "sampleP201M2",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
        "sampleP204M2"), vars=colnames(beta1))

betaM2=cbind(betaM2,beta1[,M2C])
betaM2=as.data.frame(t(betaM2))
colnames(betaM2)=colnames(beta1)

loess.linesM2=NULL

for (i in 1:nrow(betaM2)) {
  loess.linesM2=c(loess.linesM2,betaM2[i,])
}

loess.linesM2=as.data.frame(loess.linesM2)
loess.linesM2=t(loess.linesM2)

TPM2=c(rep(5,length(loess.linesM2)))

DEM2=c(rep("I",35*381),rep("C",34*381))


########
## M3 ##
########

M3=matches(c("M3"), vars=colnames(beta1))

betaM3=beta1[,M3]
rownames(betaM3)=rownames(beta1)
View(betaM3)

M3R=matches(c("sample3R451M3",
        "sample3R452M3",
        "sample3R455M3",
        "sample508M3",
        "sample509M3",
        "sample519M3",
        "sample526M3",
        "sample527M3",
        "sample540M3",
        "sample578M3",
        "sample583M3",
        "sample649M3",
        "sample663M3",
        "sample665M3",
        "sample676M3",
        "sample678M3",
        "sample694M3",
        "sample695M3",
        "sample794M3",
        "sample795M3",
        "sample801M3",
        "sample807M3",
        "sample809M3",
        "sample830M3",
        "sample833M3",
        "sample837M3",
        "sample851M3",
        "sample880M3",
        "sample884M3",
        "sample885M3",
        "sample891M3",
        "sample897M3",
        "sample907M3",
        "sample934M3",
        "sample935M3",
        "sampleP101M3",
        "sampleP201M3",
        "sampleP204M3"), vars=colnames(betaM3))

betaM3=betaM3[,-M3R]
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
M3C=matches(c("sample3R451M3",
        "sample3R452M3",
        "sample3R455M3",
        "sample508M3",
        "sample509M3",
        "sample519M3",
        "sample526M3",
        "sample527M3",
        "sample540M3",
        "sample578M3",
        "sample583M3",
        "sample649M3",
        "sample663M3",
        "sample665M3",
        "sample676M3",
        "sample678M3",
        "sample694M3",
        "sample695M3",
        "sample794M3",
        "sample795M3",
        "sample801M3",
        "sample807M3",
        "sample809M3",
        "sample830M3",
        "sample833M3",
        "sample837M3",
        "sample851M3",
        "sample880M3",
        "sample884M3",
        "sample885M3",
        "sample891M3",
        "sample897M3",
        "sample907M3",
        "sample934M3",
        "sample935M3",
        "sampleP101M3",
        "sampleP201M3",
        "sampleP204M3"), vars=colnames(beta1))


betaM3=cbind(betaM3,beta1[,M3C])
betaM3=as.data.frame(t(betaM3))
colnames(betaM3)=colnames(beta1)

loess.linesM3=NULL

for (i in 1:nrow(betaM3)) {
  loess.linesM3=c(loess.linesM3,betaM3[i,])
}

loess.linesM3=as.data.frame(loess.linesM3)
loess.linesM3=t(loess.linesM3)

TPM3=c(rep(6,length(loess.linesM3)))

DEM3=c(rep("I",35*381),rep("C",35*381))

########
## M4 ##
########

M4=matches(c("M4"), vars=colnames(beta1))

betaM4=beta1[,M4]
rownames(betaM4)=rownames(beta1)
View(betaM4)

M4C=matches(c("sample3R451M4",
        "sample3R452M4",
        "sample3R455M4",
        "sample508M4",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
        "sample509M4",
        "sample519M4",
        "sample526M4",
        "sample527M4",
        "sample540M4",
        "sample578M4",
        "sample583M4",
        "sample649M4",
        "sample663M4",
        "sample665M4",
        "sample676M4",
        "sample678M4",
        "sample694M4",
        "sample695M4",
        "sample794M4",
        "sample795M4",
        "sample801M4",
        "sample807M4",
        "sample809M4",
        "sample830M4",
        "sample833M4",
        "sample837M4",
        "sample851M4",
        "sample880M4",
        "sample884M4",
        "sample885M4",
        "sample891M4",
        "sample897M4",
        "sample907M4",
        "sample934M4",
        "sample935M4",
        "sampleP101M4",
        "sampleP201M4",
        "sampleP204M4"), vars=colnames(beta1))


betaM4=cbind(betaM4,beta1[,M4C])
betaM4=as.data.frame(t(betaM4))
colnames(betaM4)=colnames(beta1)

loess.linesM4=NULL

for (i in 1:nrow(betaM4)) {
  loess.linesM4=c(loess.linesM4,betaM4[i,])
}

loess.linesM4=as.data.frame(loess.linesM4)
loess.linesM4=t(loess.linesM4)

TPM4=c(rep(7,length(loess.linesM4)))

DEM4=c(rep("I",18*381),rep("C",33*381))


########
## M5 ##
########

M5=matches(c("M5"), vars=colnames(beta1))

betaM5=beta1[,M5]
rownames(betaM5)=rownames(beta1)
View(betaM5)

M5C=matches(c("sample3R451M5",
        "sample3R452M5",
        "sample3R455M5",
        "sample508M5",
        "sample509M5",
        "sample519M5",
        "sample526M5",
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
          "sample527M5",
          "sample540M5",
          "sample578M5",
          "sample583M5",
          "sample649M5",
          "sample663M5",
          "sample665M5",
          "sample676M5",
          "sample678M5",
          "sample694M5",
          "sample695M5",
          "sample794M5",
          "sample795M5",
          "sample801M5",
          "sample807M5",
          "sample809M5",
          "sample830M5",
          "sample833M5",
          "sample837M5",
          "sample851M5",
          "sample880M5",
          "sample884M5",
          "sample885M5",
          "sample891M5",
          "sample897M5",
          "sample907M5",
          "sample934M5",
          "sample935M5",
          "sampleP101M5",
          "sampleP201M5",
          "sampleP204M5"), vars=colnames(beta1))


betaM5=cbind(betaM5,beta1[,M5C])
betaM5=as.data.frame(t(betaM5))
colnames(betaM5)=colnames(beta1)
loess.linesM5=NULL

for (i in 1:nrow(betaM5)) {
  loess.linesM5=c(loess.linesM5,betaM5[i,])
}

loess.linesM5=as.data.frame(loess.linesM5)
loess.linesM5=t(loess.linesM5)

TPM5=c(rep(8,length(loess.linesM5)))

DEM5=c(rep("I",16*381),rep("C",36*381))

# Data frame for weighted unifrac in each timepoint (TP)

#### TP 1 ####
BC1I=betaM0[matches("M0", vars = rownames(betaM0)),matches("M0", vars = rownames(beta1))]
BC1I=BC1I[,-M0R]
colnames(BC1I)=rownames(BC1I)
BC1I[upper.tri(BC1I)] <- NA
BC1I[BC1I==0]=NA

BC1WI=NULL

for(i in 1:nrow(BC1I)){
  BC1WI=c(BC1WI,BC1I[i,])
}

BC1WI=as.data.frame(BC1WI)

F1=1;Fn=0

for(i in 1:nrow(BC1I)){
  for(j in 1:ncol(BC1I)){
    Fn=Fn+F1
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
    colnames(BC1WI)[Fn]=paste(rownames(BC1I)[i],"/",colnames(BC1I)[j])
  }
}

colnames(BC1WI)=sub("M0","",colnames(BC1WI))
colnames(BC1WI)=sub("M0","",colnames(BC1WI))

BC1C=beta1[M0C,M0C]
BC1C[upper.tri(BC1C)] <- NA
BC1C[BC1C==0]=NA
rownames(BC1C)=colnames(BC1C)

BC1WC=NULL

for(i in 1:nrow(BC1C)){
  BC1WC=c(BC1WC,BC1C[i,])
}

BC1WC=as.data.frame(BC1WC)

F1=1;Fn=0

for(i in 1:nrow(BC1C)){
  for(j in 1:nrow(BC1C)){
    Fn=Fn+F1
    colnames(BC1WC)[Fn]=paste(rownames(BC1C)[i],"/",colnames(BC1C)[j])
  }
}

colnames(BC1WC)=sub("M0","",colnames(BC1WC))
colnames(BC1WC)=sub("M0","",colnames(BC1WC))

#### TP 2 ####
BC2I=betaM1[matches("M1", vars = rownames(betaM1)),matches("M1", vars = rownames(beta1))]
BC2I=BC2I[,-M1R]
colnames(BC2I)=rownames(BC2I)
BC2I[upper.tri(BC2I)] <- NA
BC2I[BC2I==0]=NA

BC2WI=NULL

for(i in 1:nrow(BC2I)){
  BC2WI=c(BC2WI,BC2I[i,])
}

BC2WI=as.data.frame(BC2WI)

F1=1;Fn=0

for(i in 1:nrow(BC2I)){
  for(j in 1:nrow(BC2I)){
    Fn=Fn+F1
    colnames(BC2WI)[Fn]=paste(rownames(BC2I)[i],"/",colnames(BC2I)[j])
  }
}

colnames(BC2WI)=sub("M1","",colnames(BC2WI))
colnames(BC2WI)=sub("M1","",colnames(BC2WI))

BC2C=beta1[M1C,M1C]
BC2C[upper.tri(BC2C)] <- NA
BC2C[BC2C==0]=NA
rownames(BC2C)=colnames(BC2C)

BC2WC=NULL
for(i in 1:nrow(BC2C)){
  BC2WC=c(BC2WC,BC2C[i,])
}

BC2WC=as.data.frame(BC2WC)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
F1=1;Fn=0

for(i in 1:nrow(BC2C)){
  for(j in 1:nrow(BC2C)){
    Fn=Fn+F1
    colnames(BC2WC)[Fn]=paste(rownames(BC2C)[i],"/",colnames(BC2C)[j])
  }
}

colnames(BC2WC)=sub("M1","",colnames(BC2WC))
colnames(BC2WC)=sub("M1","",colnames(BC2WC))

#### TP 3 ####
BC3I=betaM2[matches("M2", vars = rownames(betaM2)),matches("M2", vars = rownames(beta1))]
BC3I=BC3I[,-M2R]
BC3I[upper.tri(BC3I)] <- NA
BC3I[BC3I==0]=NA
colnames(BC3I)=rownames(BC3I)

BC3WI=NULL

for(i in 1:nrow(BC3I)){
  BC3WI=c(BC3WI,BC3I[i,])
}

BC3WI=as.data.frame(BC3WI)

F1=1;Fn=0

for(i in 1:nrow(BC3I)){
  for(j in 1:nrow(BC3I)){
    Fn=Fn+F1
    colnames(BC3WI)[Fn]=paste(rownames(BC3I)[i],"/",colnames(BC3I)[j])
  }
}

colnames(BC3WI)=sub("M2","",colnames(BC3WI))
colnames(BC3WI)=sub("M2","",colnames(BC3WI))

BC3C=beta1[M2C,M2C]
BC3C[upper.tri(BC3C)] <- NA
BC3C[BC3C==0]=NA
rownames(BC3C)=colnames(BC3C)

BC3WC=NULL
for(i in 1:nrow(BC3C)){
  BC3WC=c(BC3WC,BC3C[i,])
}

BC3WC=as.data.frame(BC3WC)

F1=1;Fn=0

for(i in 1:nrow(BC3C)){
  for(j in 1:nrow(BC3C)){
    Fn=Fn+F1
    colnames(BC3WC)[Fn]=paste(rownames(BC3C)[i],"/",colnames(BC3C)[j])
  }
}

colnames(BC3WC)=sub("M2","",colnames(BC3WC))
colnames(BC3WC)=sub("M2","",colnames(BC3WC))

#### TP 4 ####
BC4I=betaM3[matches("M3", vars = rownames(betaM3)),matches("M3", vars = rownames(beta1))]
BC4I=BC4I[,-M3R]
BC4I[upper.tri(BC4I)] <- NA
BC4I[BC4I==0]=NA
colnames(BC4I)=rownames(BC4I)

BC4WI=NULL
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
for(i in 1:nrow(BC4I)){
  BC4WI=c(BC4WI,BC4I[i,])
}

BC4WI=as.data.frame(BC4WI)

F1=1;Fn=0

for(i in 1:nrow(BC4I)){
  for(j in 1:nrow(BC4I)){
    Fn=Fn+F1
    colnames(BC4WI)[Fn]=paste(rownames(BC4I)[i],"/",colnames(BC4I)[j])
  }
}

colnames(BC4WI)=sub("M3","",colnames(BC4WI))
colnames(BC4WI)=sub("M3","",colnames(BC4WI))

BC4C=beta1[M3C,M3C]
BC4C[upper.tri(BC4C)] <- NA
BC4C[BC4C==0]=NA
rownames(BC4C)=colnames(BC4C)

BC4WC=NULL

for(i in 1:nrow(BC4C)){
  BC4WC=c(BC4WC,BC4C[i,])
}

BC4WC=as.data.frame(BC4WC)

F1=1;Fn=0

for(i in 1:nrow(BC4C)){
  for(j in 1:nrow(BC4C)){
    Fn=Fn+F1
    colnames(BC4WC)[Fn]=paste(rownames(BC4C)[i],"/",colnames(BC4C)[j])
  }
}

colnames(BC4WC)=sub("M3","",colnames(BC4WC))
colnames(BC4WC)=sub("M3","",colnames(BC4WC))

#### TP 5 ####
BC5I=beta1[M4,M4]
BC5I[upper.tri(BC5I)] <- NA
BC5I[BC5I==0]=NA
rownames(BC5I)=colnames(BC5I)

BC5WI=NULL

for(i in 1:nrow(BC5I)){
  BC5WI=c(BC5WI,BC5I[i,])
}

BC5WI=as.data.frame(BC5WI)

F1=1;Fn=0

for(i in 1:nrow(BC5I)){
  for(j in 1:nrow(BC5I)){
    Fn=Fn+F1
    colnames(BC5WI)[Fn]=paste(rownames(BC5I)[i],"/",colnames(BC5I)[j])
  }
}

colnames(BC5WI)=sub("M4","",colnames(BC5WI))
colnames(BC5WI)=sub("M4","",colnames(BC5WI))

BC5C=beta1[M4C,M4C]
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
BC5C[upper.tri(BC5C)] <- NA
BC5C[BC5C==0]=NA
rownames(BC5C)=colnames(BC5C)

BC5WC=NULL

for(i in 1:nrow(BC5C)){
  BC5WC=c(BC5WC,BC5C[i,])
}

BC5WC=as.data.frame(BC5WC)

F1=1;Fn=0

for(i in 1:nrow(BC5C)){
  for(j in 1:nrow(BC5C)){
    Fn=Fn+F1
    colnames(BC5WC)[Fn]=paste(rownames(BC5C)[i],"/",colnames(BC5C)[j])
  }
}

colnames(BC5WC)=sub("M4","",colnames(BC5WC))
colnames(BC5WC)=sub("M4","",colnames(BC5WC))

#### TP 6 ####
BC6I=beta1[M5,M5]
BC6I[upper.tri(BC6I)] <- NA
BC6I[BC6I==0]=NA
rownames(BC6I)=colnames(BC6I)

BC6WI=NULL

for(i in 1:nrow(BC6I)){
  BC6WI=c(BC6WI,BC6I[i,])
}

BC6WI=as.data.frame(BC6WI)

F1=1;Fn=0

for(i in 1:nrow(BC6I)){
  for(j in 1:nrow(BC6I)){
    Fn=Fn+F1
    colnames(BC6WI)[Fn]=paste(rownames(BC6I)[i],"/",colnames(BC6I)[j])
  }
}

colnames(BC6WI)=sub("M5","",colnames(BC6WI))
colnames(BC6WI)=sub("M5","",colnames(BC6WI))

BC6C=beta1[M5C,M5C]
BC6C[upper.tri(BC6C)] <- NA
BC6C[BC6C==0]=NA
rownames(BC6C)=colnames(BC6C)

BC6WC=NULL

for(i in 1:nrow(BC6C)){
  BC6WC=c(BC6WC,BC6C[i,])
}

BC6WC=as.data.frame(BC6WC)

F1=1;Fn=0

for(i in 1:nrow(BC6C)){
  for(j in 1:nrow(BC6C)){
    Fn=Fn+F1
    colnames(BC6WC)[Fn]=paste(rownames(BC6C)[i],"/",colnames(BC6C)[j])
  }
}
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
colnames(BC6WC)=sub("M5","",colnames(BC6WC))
colnames(BC6WC)=sub("M5","",colnames(BC6WC))

#### all TPs #####
BCig=cbind(BC1WI,BC1WC,BC2WI,BC2WC,BC3WI,BC3WC,BC4WI,BC4WC,BC5WI,BC5WC,BC6WI,BC6WC)
BCig=t(BCig)
BCig=as.data.frame(BCig)
BCig$TimePoint=c(rep(3,(ncol(BC1WC)+ncol(BC1WI))),rep(4,(ncol(BC2WC)+ncol(BC2WI))),rep(5,(ncol(BC3WC)+ncol(BC3WI))),rep(
6,(ncol(BC4WC)+ncol(BC4WI))),rep(7,(ncol(BC5WC)+ncol(BC5WI))),rep(8,(ncol(BC6WC)+ncol(BC6WI))))
BCig$Grupo=c(rep("PR",ncol(BC1WI)),rep("Control",ncol(BC1WC)),rep("PR",ncol(BC2WI)),rep("Control",ncol(BC2WC)),rep("PR",nc
ol(BC3WI)),rep("Control",ncol(BC3WC)),rep("PR",ncol(BC4WI)),rep("Control",ncol(BC4WC)),rep("PR",ncol(BC5WI)),rep("Control",n
col(BC5WC)),rep("PR",ncol(BC6WI)),rep("Control",ncol(BC6WC)))
BCig$id=rownames(BCig)
BCig$TP=as.factor(BCig$TimePoint)
BCig$Grupo=as.factor(BCig$Grupo)
colnames(BCig)=c("BC","TimePoint","Grupo","id","TP")

BCig=BCig[complete.cases(BCig),]

library(tools)
BCig$id=file_path_sans_ext(BCig$id)

BCig$Baseline=NULL
for(j in 1:1225){
  for (i in 1:nrow(BCig)){
    if(BCig$id[i]==BCig$id[j]){
      BCig$Baseline[i]=BCig$BC[j]
    }
  }
}

BCig$Baseline[round(BCig$Baseline,7)==0.2510891]=NA #change for the first value of the matrix
BCig$Baseline[BCig$id=="sample3R493...sample3R506"]=0.2510891 #change for the first value of the matrix


#global rate of change

BCig$Norm_BC=NULL

for(i in 1:nrow(BCig)){
  BCig$Norm_BC[i]=BCig$BC[i]-BCig$Baseline[i]
}

BCig$Norm_BC[BCig$TP=="3"]=0
BCig$TimePoint=BCig$TimePoint-2

#experimental design
library(ggthemes)
ggplot(data=BCig, aes(TimePoint,Norm_BC))+
  geom_smooth(method="loess", aes(color=Grupo), se=T)+
  ylab("Weighted Unifrac (Global Rate of Change)")+
  labs(colour="Group")+
  theme_classic()+
  scale_x_continuous(labels=as.character(BCig$TimePoint),breaks=BCig$TimePoint)

library(lme4)
modelGRC1 <- lmer(Norm_BC ~ Grupo * TP + (1|id), data = BCig)
LME.GCR=Anova(modelGRC1, test.statistic = "F", type = "III")
LME.GCR

### quality assessment ###
AIC(modelGRC1)

library(MuMIn)
r.squaredGLMM(modelGRC1)

res=residuals(modelGRC1)
mean(res)

qqnorm(res, datax = TRUE)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
qqline(res, datax = TRUE)

#contrast analyses
library(gmodels)
library(lsmeans)
ref3<-lsmeans(modelGRC1 ,c("Grupo","TP"))
ref3 #check the combination of factors for contrasts
ContrastVec <- list(M3=c(1,-1,0,0,0,0,0,0,0,0,0,0),M4=c(0,0,1,-1,0,0,0,0,0,0,0,0), M5=c(0,0,0,0,1,-1,0,0,0,0,0,0),M6=c(0,0,0,0,0,0,1,-
1,0,0,0,0),M7=c(0,0,0,0,0,0,0,0,1,-1,0,0),M8=c(0,0,0,0,0,0,0,0,0,0,1,-1))
summary(contrast(ref3, ContrastVec), adjust = "bonferroni")


#descriptive analysis' table
M_WIG.G=aggregate(Norm_BC~ Grupo+TimePoint, data = BCig, FUN= "mean" )
SD_WIG.G=aggregate(Norm_BC~ Grupo+TimePoint, data = BCig, FUN= "sd" )
SD_WIG.G$BC=as.numeric(SD_WIG.G$Norm_BC)
Table7=M_WIG.G
Table7$TimePoint=factor(rep(c("M3","M4","M5","M6","M7","M8"), each=2), levels = c("M3","M4","M5","M6","M7","M8"))
Table7$Norm_BC=round(as.numeric(Table7$Norm_BC),2)
View(Table7)

for(i in 1:12){
  Table7$Norm_BC[i]=paste(Table7$Norm_BC[i],"\u00B1", round(SD_WIG.G$Norm_BC[i],2))}

colnames(Table7)=c("Grupo","Timepoint","Mean \u00B1 SD")

ref3=as.data.frame(ref3)

Table7$lsmean=NULL
for(i in 1:12){
  Table7$lsmean[i]=paste(round(ref3[i,3],2),"\u00B1",round(ref3[i,4],3))
}

Table7=Table7[,c(2,1,3,4)]

library(kableExtra)
kbl(Table7) %>%
  kable_classic(full_width=F) %>%
  column_spec(1, bold = T) %>%
  row_spec(c(0),bold = T)%>%
  collapse_rows(columns = 1, valign = "middle")

#library(xlsx)
#write.xlsx(Table7,"Tabelas descritivas.xlsx",  sheetName ="Wheighted_U_IG_BL",append=TRUE)
#write.xlsx(Table7,"Tabelas descritivas.xlsx",  sheetName ="BC_IG_BL",append=TRUE)

ggline(BCig, x="TP",y="Norm_BC",add = c("mean_se"), color = "Grupo")+ylab("Weighted Unifrac")#Mean plot

ggline(ref3, x="TP",y="lsmean",add = c("mean_se"), color = "Grupo")+ylab("Weighted Unifrac")#Mean plot

#LSMEANS graphs

ref1=as.data.frame(ref1)
colnames(ref1)[1]="Group"
lsmeans1=ref1[,c(1,2,3)]

library(ggplot2)
library(ggpubr)
ggline(ref1, x="TP",y="lsmean", color = "Group")+ylab("lsmeans Weighted Unifrac")

ref3=as.data.frame(ref3)
colnames(ref3)[1]="Group"
lsmeans3=ref3[,c(1,2,3)]


library(ggplot2)
library(ggpubr)
ggline(ref3, x="TP",y="lsmean", color = "Group")+ylab("lsmeans GRC Weighted Unifrac")+labs(colour="Group")

# Effect size
library(effectsize)
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
library(lme4)
library(car)
library(lmerTest)
model2.1=lmer(BC ~ Grupo * TP + Baseline + (1|id), data = BCig)
summary(model2.1)
summary(model2)
anova(model2.1)#Kenward-Roger (Kenward & Roger, 1997) and Satterthwaite (1941) approach to calculate residuals degrees of freedom
summary(model2)
F_to_eta2(14,5,4356,ci=0.95)
```

Linear Mixed effect models, contrast analysis and Loess lines plots (e.g. longitudinal dynamic of Dialister between R and NR to mBorg)

```
library(readxl)
longitudinal_data <- read_excel("All_OTUs_longitudinal_compl.xlsx")

#removal of genera that were present in <20% of the samples
OTU_selected=NULL

n=1
for(i in 8:ncol(All_OTUs_longitudinal_compl)){
 if(sum(Dados_clean[,i]>0)>=83){
  OTU_selected[n]=colnames(All_OTUs_longitudinal_compl)[i]
  n = n + 1
 }
}

Dados_clean_filtered=Dados_clean[,OTU_selected]
Dados_clean_filtered=as.data.frame(cbind(Dados_clean_filtered,Dados_clean[,c(7:8)],Dados_clean[,1:6],Dados_clean[,204]))
#R e NR
Dados_clean_filtered[,81]=as.factor(Dados_clean_filtered[,81])
Dados_clean_filtered[,82]=as.factor(Dados_clean_filtered[,82])
Dados_clean_filtered[,83]=as.factor(Dados_clean_filtered[,83])
Dados_clean_filtered[,84]=as.factor(Dados_clean_filtered[,84])
Dados_clean_filtered[,85]=as.factor(Dados_clean_filtered[,85])
Dados_clean_filtered[,86]=as.factor(Dados_clean_filtered[,86])

Dados_PR=subset(Dados_clean_filtered,Group=="PR")

#linear mixed effect models between R and NR to mBorg

Resultados_modelos_mBorg=data.frame(OTU=NULL,Statistic=NULL,Pvalue=NULL,Assumptions=NULL)

library(lme4)
library(car)
n=1
for(i in 1:79){
 modelo1 <- lmer(Dados_PR_raw[,i] ~ TP*RmBorg  + (1|OTU_ID), data = Dados_PR)
 LME=Anova(modelo1, test.statistic = "F", type = "III")
 if(LME[4,4]<0.05){
  Resultados_modelos_mBorg[n,1]=colnames(Dados_PR)[i]
  Resultados_modelos_mBorg[n,2]=LME[4,1]
  Resultados_modelos_mBorg[n,3]=LME[4,4]
  n=n+1
 }
}
write.xlsx(Resultados_modelos_mBorg,"OTUs_ASVs_long.xlsx", append = TRUE, sheetName = "mBorg")

#validation of assumptions
modelo_dia_mborg <- lmer(`k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister`~ TP*RmBorg  + (1|OTU_ID), data = Dados_PR)
res=residuals(modelo_dia_mborg)
qqnorm(res)
qqline(res)
plot(modelo1)

#contrast analysis
ref1<-lsmeans(modelo_sch_bmsr,c("RmBorg","TP"))
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
ref1 #check the combination of factors for contrasts
ContrastVec <- list(M1=c(1,-1,0,0,0,0,0,0,0,0,0,0),M2=c(0,0,1,-1,0,0,0,0,0,0,0,0), M3=c(0,0,0,0,1,-1,0,0,0,0,0,0),M4=c(0,0,0,0,0,0,1,-
1,0,0,0,0),M5=c(0,0,0,0,0,0,0,0,1,-1,0,0),M6=c(0,0,0,0,0,0,0,0,0,0,1,-1))
summary(contrast(ref1, ContrastVec), adjust = "bonferroni")

#loess lines plot
ggplot(data=Dados_,
aes(TP,Dados_PR$`k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister`))+
  geom_smooth(method="loess", aes(fill=RmBORG,color= RmBORG, group= RmBORG), alpha=0.3)+
  ylab("Abundance")+
  labs(colour='Response to mBORG', fill='Response to mBORG')+
  labs(colour="Response to mBORG ")+
  theme_classic()+
  scale_x_continuous(labels=as.character(Dados_clean_filtered$TP),breaks=Dados_clean_filteres$TP)+
  ggtitle("Dialister")
```

Repeated-measures correlation (e.g. correlations between the log10 of inflammatory cytokines and arcsine square root transformed genera relative abundances in NR to exercise capacity)

```
###NR_6MWT####

library(readxl)
longitudinal_data <- read_excel("log10_longitudinal_NR6MWT_l6.xlsx")

longitudinal_data[,5:198]=apply(longitudinal_data[,5:198],2,as.numeric)#selecionar colunas com os dados numéricos

longitudinal_data$`#OTU ID`=gsub("\\M.*","",longitudinal_data$`#OTU ID`)#retirar os M1, M2, M3, M4
longitudinal_data

#removal of genera that were present in <20% of the samples
OTU_selected=NULL

n=1
for(i in 18:ncol(longitudinal_data)){
  if(sum(longitudinal_data[,i]>0)>=7)
    OTU_selected[n]=colnames(longitudinal_data)[i]
    n = n + 1
  }
}
longitudinal_data_d=longitudinal_data[,OTU_selected]
#transformation of relative abundances with arcsine square root transformation
longitudinal_data_e=asin(sqrt(longitudinal_data_d))
longitudinal_data_f=as.data.frame(cbind(longitudinal_data_e,longitudinal_data[,1:17
sampleID=NULL

#Sample ID
i=1
n=1
j=1
while(i<=nrow(longitudinal_data_f)){
  if(longitudinal_data_f$`#OTU ID`[i]==longitudinal_data_f$`#OTU ID`[j]){
    sampleID[i]=n
    i=i+1
  } else if (longitudinal_data_f$`#OTU ID`[i]!=longitudinal_data_f$`#OTU ID`[j]){
    n = n + 1
    j=i
  }
}
longitudinal_data_f$sampleID=sampleID



Matriz_corr=data.frame(NULL)
Pvalues=data.frame(NULL)

library(rmcorr)
for(i in 68:80){#columns with cytokines
  for(n in 1:63){#columns with OTUs/ASVs
    r=rmcorr(longitudinal_data_f$sampleID,longitudinal_data_f[,i],longitudinal_data_f[,n],dataset = longitudinal_data_f)
    Matriz_corr[i-67,n]=r$r
    Pvalues[i-67,n]=r$p
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
  }
}

rownames(Matriz_corr)=colnames(longitudinal_data_f[68:80])#Colunas das citocinas
rownames(Matriz_corr)=gsub("\\ .*","",colnames(longitudinal_data_f[68:80]))#Colunas das citocinas


#Reducing genera taxonomic annotation
for(i in 1:length(colnames(Matriz_corr))){
  if(grepl("g__",colnames(longitudinal_data_f)[i])&
!grepl("f__;",colnames(longitudinal_data_f)[i])|!grepl("g__",colnames(longitudinal_data_f)[i])){
    colnames(Matriz_corr)[i]=paste(gsub(".*;f__",  "",  colnames(longitudinal_data_f)[i]),  gsub("g*.",  "",
colnames(longitudinal_data_f)[i]))
  } else if (grepl("f__;",colnames(longitudinal_data_f)[i])& !grepl("o__;",colnames(longitudinal_data_f)[i])){
    colnames(Matriz_corr)[i]=paste(gsub(".*;o__",  "",  colnames(longitudinal_data_f)[i]),  gsub("f*.",  "",
colnames(longitudinal_data_f)[i]))
  } else if (grepl("o__;",colnames(longitudinal_data_f)[i])& !grepl("c__;",colnames(longitudinal_data_f)[i])) {
    colnames(Matriz_corr)[i]=paste(gsub(".*;c__",  "",  colnames(longitudinal_data_f)[i]),  gsub("o*.",  "",
colnames(longitudinal_data_f)[i]))
  }
}
colnames(Matriz_corr)
rownames(Pvalues)=colnames(longitudinal_data_f[68:80])#Colunas das citocinas
rownames(Pvalues)=gsub("\\ .*","",colnames(longitudinal_data_f[68:80]))
colnames(Pvalues)=paste(gsub(".*_", "", colnames(longitudinal_data_f)[1:63]) )#Colunas das OTUs ASVs
#export correlation matrix
write.xlsx(Matriz_corr,"Matrix_log10_Citocinas_arcsin_OTUs_NR_6MWTl6.xlsx")

#assessment of assumptions validation
corr_sign=data.frame(Correlation=NULL, P=NULL, Coefficient=NULL, Norm_Validation=NULL)
l=0
for(i in 1:63) {#colunas das ASVs OTUs na matrix de corr
  for(n in 1:13){#linhas das citocinas na matrix de corr
    if(Pvalues[n,i]<0.05){
      l=l+1
      corr_sign[l,1]=paste(colnames(Matriz_corr)[i],"VS",rownames(Matriz_corr)[n])
      corr_sign[l,2]=Pvalues[n,i]
      corr_sign[l,3]=Matriz_corr[n,i]

    }
  }

}
colnames(corr_sign)=c("Correlation","P-value", "Coefficient")
normal_test=as.data.frame(NULL)
library(rmcorr)
for(i in 68:80){#Colunas das citocinas
  for(n in 1:63){#Colunas das ASVs OTUs
    r=rmcorr(longitudinal_data_f$sampleID,longitudinal_data_f[,i],longitudinal_data_f[,n],dataset = longitudinal_data_f)
    res=residuals(r$model)
    shap=shapiro.test(res)[2]
    normal_test[i-67,n]=shap#Mudar número para a primeira coluna das citocinas menos 1.
  }
}

#normally distributed residuals
l=1
for(i in 1:63) {#columns with OTUs/ASVs
  for(n in 1:13){#rows with cytokines
    l=l+1
    if(normal_test[n,i]<0.05){
      corr_sign[l,4]="Not Validated"
    } else{
      corr_sign[l,4]="Validated"
    }
  }

}
corr_sign=corr_sign[complete.cases(corr_sign),]
colnames(corr_sign)[4]="Residual Normal Distribution"
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
library(xlsx)
write.xlsx(corr_sign,"log10_Citocinas_OTUs_NR_6MWTl6.xlsx")
```

Correlation network plots (e.g. network based on the significant correlations between the log10 of inflammatory cytokines and arcsine square root transformed genera relative abundances in NR to mBorg)

```
library(readxl)
NR_BMSD <- read_excel("Matrix_log10_Citocinas_arcsin_OTUs_NR_mBorg.xlsx")
library(tidyr)

OTUs=c(rep(c('ASV_streptococcus','Lachnoanaerobaculum','Butyrivibrio','Catonella','Lachnospiraceae',  'Lautropia'),each  =
13))#Colocar o nome das bactérias que se quer representar
citocinas = c(rep(NR_BMSD$...1,6))

corr=c(NR_mBorg$str_sp,
     NR_mBorg$`Lachnospiraceae;g__Lachnoanaerobaculum`,
     NR_mBorg$`Lachnospiraceae;g__Butyrivibrio`,
     NR_mBorg$`Lachnospiraceae;g__Catonella`,
     NR_mBorg$`Lachnospiraceae;__`,
     NR_mBorg$`Burkholderiaceae;g__Lautropia`)

data=data.frame(OTU = OTUs, Citokynes = citocinas, Correlation = corr)#all data
data
data_sign = data[c(2,17,18,20,29,41,42,45,47,55,58,60,61,65,68,73),]#select only the correlations validating the assumptions
data_sign <- data_sign[order(data_sign$Correlation),]
data_sign

nodes_sign <- data.frame(
 name=c('ASV_streptococcus','Lachnoanaerobaculum','Butyrivibrio','Catonella','Lachnospiraceae', 'Lautropia',
       NR_mBorg$...1[c(2,3,4,5,6,7,8,9,13)]), #cytokines
 carac=c(rep("Bacteria",6), rep('Citokyne',9))
)


#nodes <- data.frame(
 # name=c('Lautropia','Rothia','Gemellaceae',Matrix_log10_Citocinas_arcsin_OTUs_NR_6MWTl6$...1),
  #carac=c(rep("Bacteria",3), rep('Citokyne',13))
#)

library(igraph)
# Turn it into igraph object
network <- graph_from_data_frame(d=data_sign, vertices=nodes_sign, directed=F)
V(network)

# Make a palette of colors
vcolrs<-c('goldenrod1','goldenrod1','goldenrod1','goldenrod1','goldenrod1','red','papayawhip',
'papayawhip','papayawhip','papayawhip','papayawhip','papayawhip','papayawhip','papayawhip')
ecolrs <- ifelse( E(network)$Correlation >0, "firebrick3", "deepskyblue3")
deg <- degree(network, mode="all")*1.5

# Make the plot
plot(network, vertex.color=vcolrs, edge.color=ecolrs,
    vertex.label.color="black",
    vertex.frame.color="#ffffff",
    vertex.shape="circle",
    vertex.label.family="Arial",
    vertex.label.cex=1,
    edge.width=as.integer(abs(E(network)$Correlation)*10),
    vertex.size=deg*10,
    edge.label.cex=0.8)

#open a graphical interface to re-arrange the plot
tkplot(network, vertex.color=vcolrs, edge.color=ecolrs,
     vertex.label.color="black",
     vertex.frame.color="#ffffff",
     vertex.shape="circle",
     vertex.label.family="Arial",
     vertex.label.cex=1,
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

```
    edge.width=as.integer(abs(E(network)$Correlation)*10),
    vertex.size=deg*10,
    edge.label.cex=0.8)

write_graph(network,  "NR_BMSD_network.gml", format = "gml")
```

Pitman-Morgan paired variance test (e.g. Variance of Il-1β in NR to exercise capacity)

```
library(readxl)
var_6mwt <- read_excel("var_6mwt.xlsx")

library(PairedData)
Var.test(var_6mwt$Il1bM1NR, var_6mwt$IL1bM2NR,paired=TRUE)
#p=0.01
Var.test(var_6mwt$IL1bM2NR, var_6mwt$IL1bM4NR,paired=TRUE)
#p=0.007
Var.test(var_6mwt$IL1bM1NR, var_6mwt$IL1bM4NR,paired=TRUE)
#p=0.05

#adjusting for multiple comparisons
a<-c(0.01, 0.007, 0.053)
p.adjust(a, method = "bonferroni", n = 3) #0.030 0.021 0.159
```

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

References

1 Marques A, Jácome C, Rebelo P, *et al.* Improving access to community-based pulmonary rehabilitation: 3R protocol for real-world settings with cost-benefit analysis. *BMC Public Health* 2019;**19**:676. doi:10.1186/s12889-019-7045-1

2 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83. doi:10.1016/0021-9681(87)90171-8

3 GOLD - Global Strategy for Diagnosis, Management, and prevention of chronic obstructive pulmonary disease 2020 report. 2020. 1–141.

4 Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement | American Journal of Respiratory and Critical Care Medicine. https://www.atsjournals.org/doi/full/10.1164/rccm.201908-1590ST (accessed 28 Oct 2022).

5 Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *European Respiratory Journal* 2009;**34**:648–54. doi:10.1183/09031936.00102509

6 Jones PW, Tabberer M, Chen W-H. Creating scenarios of the impact of COPD and their relationship to COPD Assessment Test (CAT™) scores. *BMC Pulm Med* 2011;**11**:42. doi:10.1186/1471-2466-11-42

7 ATS Statement. *Am J Respir Crit Care Med* 2002;**166**:111–7. doi:10.1164/ajrccm.166.1.at1102

8 Holland AE, Spruit MA, Troosters T, *et al.* An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. *European Respiratory Journal* 2014;**44**:1428–46. doi:10.1183/09031936.00150314

9 Jones PW, Beeh KM, Chapman KR, *et al.* Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014;**189**:250–5. doi:10.1164/rccm.201310-1863PP

10 Holland AE, Hill CJ, Rasekaba T, *et al.* Updating the minimal important difference for six-minute walk distance in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil* 2010;**91**:221–5. doi:10.1016/j.apmr.2009.10.017

11 Bolyen E, Rideout JR, Dillon MR, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7. doi:10.1038/s41587-019-0209-9

12 Baeshen HA. Assessment of salivary pro inflammatory cytokines profile level in patients treated with labial and lingual fixed orthodontic appliances. *PLoS One* 2021;**16**:e0249999. doi:10.1371/journal.pone.0249999

13 Maydych V, Claus M, Watzl C, *et al.* Attention to Emotional Information Is Associated With Cytokine Responses to Psychological Stress. *Frontiers in Neuroscience* 2018;**12**.https://www.frontiersin.org/articles/10.3389/fnins.2018.00687 (accessed 25 Nov 2022).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

14 LEGENDplex™. https://www.biolegend.com/en-us/legendplex (accessed 21 Mar 2022).

15 Prism - GraphPad. https://www.graphpad.com/scientific-software/prism/ (accessed 12 Jan 2022).

16 R: The R Project for Statistical Computing. https://www.r-project.org/ (accessed 12 Jan 2022).

17 QIIME 2. https://qiime2.org/ (accessed 12 Jan 2022).

18 vsearch — QIIME 2 2020.8.0 documentation. https://docs.qiime2.org/2020.8/plugins/available/vsearch/ (accessed 30 Dec 2021).

19 deblur — QIIME 2 2020.8.0 documentation. https://docs.qiime2.org/2020.8/plugins/available/deblur/ (accessed 30 Dec 2021).

20 Amir A, McDonald D, Navas-Molina JA, *et al.* Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* 2017;**2**:e00191-16, /msys/2/2/e00191-16.atom. doi:10.1128/mSystems.00191-16

21 Callahan B, Davis NM, Ernst FGM. decontam: Identify Contaminants in Marker-gene and Metagenomics Sequencing Data. 2021. doi:10.18129/B9.bioc.decontam

22 Callahan B. benjjneb/decontam. 2021.https://github.com/benjjneb/decontam (accessed 12 Jan 2022).

23 Phylogenetic inference with q2-phylogeny — QIIME 2 2021.11.0 documentation. https://docs.qiime2.org/2021.11/tutorials/phylogeny/?highlight=phylogeny (accessed 30 Dec 2021).

24 Katoh K, Misawa K, Kuma K, *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66. doi:10.1093/nar/gkf436

25 FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0009490 (accessed 23 Jul 2019).

26 Bokulich NA, Kaehler BD, Rideout JR, *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 2018;**6**:90. doi:10.1186/s40168-018-0470-z

27 QIIME 2 Library. https://library.qiime2.org/plugins/q2-feature-classifier/3/ (accessed 30 Dec 2021).

28 F. Escapa I, Huang Y, Chen T, *et al.* Construction of habitat-specific training sets to achieve species-level assignment in 16S rRNA gene datasets. *Microbiome* 2020;**8**:65. doi:10.1186/s40168-020-00841-w

29 diversity — QIIME 2 2020.8.0 documentation. https://docs.qiime2.org/2020.8/plugins/available/diversity/ (accessed 30 Dec 2021).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

30 R: The R Stats Package. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html (accessed 12 Jan 2022).

31 Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 2001;**26**:32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

32 adonis function - RDocumentation. https://www.rdocumentation.org/packages/vegan/versions/2.4-2/topics/adonis (accessed 15 Sep 2021).

33 vegan-package: Community Ecology Package: Ordination, Diversity and... in vegan: Community Ecology Package. https://rdrr.io/cran/vegan/man/vegan-package.html (accessed 12 Jan 2022).

34 Mandal S, Van Treuren W, White RA, *et al.* Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis* 2015;**26**:27663. doi:10.3402/mehd.v26.27663

35 Segata N, Izard J, Waldron L, *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;**12**:R60. doi:10.1186/gb-2011-12-6-r60

36 Huttenhower C. Galaxy / Hutlab - Harvard. http://huttenhower.sph.harvard.edu/galaxy/ (accessed 19 Sep 2019).

37 Mandal, Siddhartha. Research - Dr. Siddhartha Mandal: ANCOM 2.0 - updated code. https://sites.google.com/site/siddharthamandal1985/research (accessed 19 Sep 2019).

38 Kaul A, Mandal S, Davidov O, *et al.* Analysis of Microbiome Data in the Presence of Excess Zeros. *Front Microbiol* 2017;**8**:2114. doi:10.3389/fmicb.2017.02114

39 Mandal S, Van Treuren W, White RA, *et al.* Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial Ecology in Health & Disease* 2015;**26**. doi:10.3402/mehd.v26.27663

40 Lin FH. User Manual for ANCOM v2.1. 2022.https://github.com/FrederickHuangLin/ANCOM (accessed 31 Jan 2022).

41 Bates D, Mächler M, Bolker B, *et al.* Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 2015;**67**:1–48. doi:10.18637/jss.v067.i01

42 PairedData.pdf. https://cran.r-project.org/web/packages/PairedData/PairedData.pdf (accessed 1 Aug 2022).

43 Bakdash JZ, Marusich LR. Repeated Measures Correlation. *Frontiers in Psychology* 2017;**8**.https://www.frontiersin.org/article/10.3389/fpsyg.2017.00456 (accessed 15 Mar 2022).

44 igraph – Network analysis software. https://igraph.org/ (accessed 22 Jul 2022).

45 Melo-Dias S, Valente C, Andrade L, *et al.* Saliva as a non-invasive specimen for COPD assessment. *Respiratory Research* 2022;**23**:16. doi:10.1186/s12931-022-01935-9

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

2924

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

STROBE Statement—Checklist of items that should be included in reports of cohort studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract **YES**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found **YES**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **YES**
Objectives	3	State specific objectives, including any prespecified hypotheses **YES**
Methods		
Study design	4	Present key elements of study design early in the paper **YES**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **YES**
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up **YES**
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed **YES (detailed description in Supplementary file)**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable **YES (detailed description in Supplementary file)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **YES (detailed description in Supplementary file)**
Bias	9	Describe any efforts to address potential sources of bias **YES (detailed description in Supplementary file)**
Study size	10	Explain how the study size was arrived at **n.a.**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **YES (detailed description in Supplementary file)**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **YES (detailed description in Supplementary file)**

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

		(*b*) Describe any methods used to examine subgroups and interactions **YES (detailed description in Supplementary file)**
		(*c*) Explain how missing data were addressed **YES (detailed description in Supplementary file)**
		(*d*) If applicable, explain how loss to follow-up was addressed **YES (detailed description in Supplementary file)**
		(*e*) Describe any sensitivity analyses **n.a.**
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed **YES**
		(b) Give reasons for non-participation at each stage **YES**
		(c) Consider use of a flow diagram **n.a.**
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **YES**
		(b) Indicate number of participants with missing data for each variable of interest **YES**
		(c) Summarise follow-up time (eg, average and total amount) **YES**
Outcome data	15*	Report numbers of outcome events or summary measures over time
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included **YES**
		(*b*) Report category boundaries when continuous variables were categorized **YES**
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period **n.a**
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses **YES**
Discussion		
Key results	18	Summarise key results with reference to study objectives **YES**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **YES**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **YES**

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Generalisability	21	Discuss the generalisability (external validity) of the study results **YES**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based **YES**

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Summary of DECONTAM results preformed with the prevalence method and a threshold of 0.5 for identification of contaminants in the dataset.

ASV code	Prevalence in samples	Score	Contaminant
f65ef8a64333fc9aa05887b2aa16f257	30	4.93566934530699E-35	TRUE
8c8598245ea67244fb156b6f31192a41	51	6.55556091875317E-22	TRUE
4bb2a60ebd87faa796298b17bcc4ebef	122	7.12476858009646E-20	TRUE
6c891abaa8f2147383dd332e601800eb	25	5.64462712512244E-19	TRUE
073d0c62d1257a783d7508c138432fc4	22	5.95502824269748E-19	TRUE
609cfe075d3af9506da2e29b0db659fe	20	8.07960447514411E-17	TRUE
20ad5b343a158e617aa2f39adbf005b6	18	1.29146710631416E-12	TRUE
e191f1f77e171985e4ce4e16cfd52089	11	1.13823341316519E-11	TRUE
39032f4b6765815d5012773e6807e4be	14	1.98170460521423E-11	TRUE
d65886bee0f292fbbf5fc64ff1bc0b24	221	1.6692749528622E-10	TRUE
fd657e81a8537d9ff329c0703a0f7f63	10	1.72962979000839E-10	TRUE
20cbb7a8b00c1670d0175f2d742e655b	9	2.52133562311872E-09	TRUE
217791e2b9f75aa3915ad540e35c3612	13	1.0669565420676E-08	TRUE
54be6aa07bc7127f2df3048332e7f0cc	6	7.05006830680515E-08	TRUE
fd2b0b0983f5c7a61e3117545344180c	12	1.19572186337503E-07	TRUE
d23fbef2f31d48eda40876cdbc49933a	194	2.32906586692203E-07	TRUE
574d164310944193d8fc13dc10346e58	45	3.99850620766581E-07	TRUE
1db319dd24afdc2890163fd4cfd776e0	7	0.0000004703898033813	TRUE
8fbb45639f32c85085e06660994f97ae	7	0.0000004703898033813	TRUE
083310805ab3a26fd0b20219868d894b	10	4.78885615151771E-07	TRUE
11f63e800d538ce61113253faf0f0363	10	4.78885615151771E-07	TRUE
e15cb616bbfc94e116f873ba1b50aabc	10	4.78885615151771E-07	TRUE
ed7a396f4b38c1b98b597e923ae53f22	5	1.04469193996803E-06	TRUE
e941a4a1ad9617d44833702de4f0ca82	5	1.04469193996803E-06	TRUE
2bbd06be7ee1e2cef688d5d3b880053f	5	1.04469193996803E-06	TRUE
fca9741d7bcdc137ad78a45a0f50c3b6	30	1.36535722961629E-06	TRUE
5785bb04044f25bd01768aa97d019b2e	8	1.79321917846439E-06	TRUE
8f09bac714c02dc9415770271349dc99	196	1.80482872811538E-06	TRUE
77c55b0f6c379835576843cf02171325	17	4.57816661895207E-06	TRUE
05592ceac1f1697c9020de02157cec66	6	5.98614890787024E-06	TRUE
1187d3f6bbc3c560b79f47cf9ad873bd	49	9.12411188077238E-06	TRUE
c7ef526f7263c4a3683d096b6748a0fb	4	0.00001505585442969 55	TRUE
98bc1ec29a4c0f9fbf70d7bfa48a48b6	4	0.00001505585442969 55	TRUE
1ea83794b96a00695927de44a90aa26d	4	0.00001505585442969 55	TRUE
d8f298eccc87fa0854033cac0ce4baca	4	0.00001505585442969 55	TRUE
9ab98702e9cf7c27aa339109ea83e689	4	0.00001505585442969 55	TRUE
57ef5006312a9662d87d57c1404c1a87	7	0.0000200061185749385	TRUE
548c3f420c3c6a1eb10f35aed1eeb0c1	7	0.0000200061185749385	TRUE
eefaf42aa3481703873ea5b975cca9bf	15	0.0000200562316647648	TRUE
6ba560a76661744bc65eed04953c1314	5	0.0000721451963282275	TRUE
7f1487c191bfa6ccc9085821aadf4324	65	0.0000897630873541179	TRUE
891019e319f5d8d088ed4f9379fd63ef	208	0.000164834348740025	TRUE
ec9d0f20c8c61eb5032d07937cf1af3c	6	0.000207385864939935	TRUE
0a738d348f5e152a2521b6eb922ef446	6	0.000207385864939935	TRUE
365bc413ed3cf82d8737939b5581c3cc	6	0.000207385864939935	TRUE
791ca5935b55de1be66b9c83d56e4167	3	0.000211212129284659	TRUE
c873733a1c472c123f8063297a02c773	3	0.000211212129284659	TRUE
42a8ef3012c07a08d6fbabaef0273cbb	3	0.000211212129284659	TRUE
a20597397ea624f1d40736bbb973dff1	3	0.000211212129284659	TRUE
ee02b9d7321d96af270ccd201ba17b6b	3	0.000211212129284659	TRUE
048633e2cfb6cd41a06beacf26f04cb6	3	0.000211212129284659	TRUE
6fadc82ce33939380ddb960da7f411c1	3	0.000211212129284659	TRUE
f075744cdedbbbec727c7e1b2037c3e6	3	0.000211212129284659	TRUE
30554407069b2367b77ee68c1982db14	3	0.000211212129284659	TRUE
b2bba50ae5022b13d0cd6641814e0713	3	0.000211212129284659	TRUE
79544aab951bcf5f54f663c1dc3f2912	3	0.000211212129284659	TRUE
b83aaa7e85c9c03fa9485e12be27116b	3	0.000211212129284659	TRUE
e58f4afd38d5d2317e576701eac632a4	3	0.000211212129284659	TRUE
8d3565638a5a89a00aa3233b2b027489	4	0.000814736808279286	TRUE
7bae376e41b6259952179d259f069c1e	4	0.000814736808279286	TRUE
5b84f9aefbea73054bcb5ebbf3904893	4	0.000814736808279286	TRUE
a8a7c72cd7f5319290396ea098c97020	4	0.000814736808279286	TRUE
e6f6be4ccfb278fe562ca7dffb543652	9	0.00153092250680292	TRUE
32cfdd0837255fb3842fc282e432ec52	9	0.00153092250680292	TRUE
e795fed13ead27da7c3bedd15b4a59c1	5	0.00196365213243004	TRUE
a5e3cfb18eb2fd335108bf5b8d60b739	5	0.00196365213243004	TRUE
d62009c3d7662c9c54d23242713f314e	2	0.00288656576689034	TRUE
d5695af3afad7256ad94ba6a38ccd4fb	2	0.00288656576689034	TRUE
e2fcc92ed4f145071525b58f2f750e39	2	0.00288656576689034	TRUE
e55f7ffdc508fc02e0ce54bfb21e6038	2	0.00288656576689034	TRUE
7fa99bd3ee34c6d339e3ff01bd88427e	2	0.00288656576689034	TRUE
0eb8e63cb8f7cb7aa58b403dc29da7de	2	0.00288656576689034	TRUE
464a52deb6e8c4e2b57d980a8d18f04c	2	0.00288656576689034	TRUE
2d34c22edce4b1f2d8a5228ad78f8ea8	2	0.00288656576689034	TRUE
6a9e6cd156ec3e7ca51f4912bfe49614	2	0.00288656576689034	TRUE
da8b2afed60ab7f62166e969e4322f11	6	0.0037851032560835	TRUE
f92c2786afd33f8c9eddd65bcb666f42	18	0.00487193302555845	TRUE
50a3b56e0b7db75ee9daccf7a751ba41	19	0.00630479812169273	TRUE
84716245f53ab2236b84538a9cdd74ed	7	0.0063823686613895	TRUE
ed2e115df9ed4a9cf8de15df330c428e	262	0.00647182724575958	TRUE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

cfeb33d30aa1b73ecca815bc65ff35f0	3	0.00844848517138625	TRUE
038cda24080464defef44077e55d614a	3	0.00844848517138625	TRUE
1772eec427522cda596e760ad3ebef61	3	0.00844848517138625	TRUE
a5d1217aa1cbc154c499796f9040d954	14	0.00977619312112795	TRUE
6628b08e13117c52417f5dfee4d9eafd	8	0.00983688431712284	TRUE
b4587918bfe7955deb4fe9932ec07c3d	4	0.0164745460842033	TRUE
2ce83ce8633190af0e5f51af6c6d7da8	4	0.0164745460842033	TRUE
31de871bad103340eaac73b2383506fe	10	0.0195459279544366	TRUE
b57744bd30af9f7a4d538d0aba27596b	11	0.0258714571922254	TRUE
d50706f84ad2be04b147f0a3b36039d3	5	0.0267556257599366	TRUE
1fa7e42cad8192ac5a0f3e2319fb1089	5	0.0267556257599366	TRUE
7b922bd11ea3f2de8a1d01bc0314871f	438	0.0311556538676718	TRUE
bc3d8f06abbe8f3820be7ab8cc2eb279	6	0.0390864277175264	TRUE
0d57a518ec5291ce0ee0643456940a90	38	0.0399684744186126	TRUE
1369aba02dd8b749b4c4c2a799d34207	13	0.0415311220541261	TRUE
107aa7e6b56274803bffd1ccff5b2ee6	14	0.050854414641935	TRUE
d353248c8a7d31bc9f8378a320b77924	14	0.050854414641935	TRUE
8bc54d575f88d09b2ba16b7f571812ae	31	0.0517381435573177	TRUE
6ba7d4f61fd8143e8aa856596c09b9dd	7	0.0532669077242673	TRUE
4908017090823380efa89d3e236c811e	33	0.0655854886123842	TRUE
dc8201773df81f21b280e07a6f320208	8	0.0691034263704471	TRUE
1ea779464fee4aab61b2adbe93db1d13	8	0.0691034263704471	TRUE
7bfaa7f90dff6c876675062e176f2b47	8	0.0691034263704471	TRUE
4dd2f7772c7e05fe793cb8d593a59dd0	8	0.0691034263704471	TRUE
642137de3ea503ff59288deb908c51fa	8	0.0691034263704471	TRUE
aeb091a4d196e216c7ed9c8cc29f36a8	8	0.0691034263704471	TRUE
67cea148133734a4183fb020f6740b4e	2	0.0769230769230769	TRUE
83e8f941ebfaf966ed1b7d81e21e43a6	2	0.0769230769230769	TRUE
403c354ef8601d5aba74b7ee78058e87	2	0.0769230769230769	TRUE
0d3f7ffbeccd21fbb5aa31e55f638069	2	0.0769230769230769	TRUE
508b4814a200ebb7843f78ebe4251e05	2	0.0769230769230769	TRUE
3caf30133a391f2f85e2365d24923375	2	0.0769230769230769	TRUE
d630cbac6a5674e8efe19b8a32751f57	2	0.0769230769230769	TRUE
65dfa12d95100cd914a50aa14ab0c74e	2	0.0769230769230769	TRUE
33c14b40a4d663250dffefdb3ce1f5b4	2	0.0769230769230769	TRUE
8e364327698c6e82d690bae5d042c737	2	0.0769230769230769	TRUE
80feb2fb070545be75f78cbd6b520d5f	2	0.0769230769230769	TRUE
8253a32e37cc65a796a92ee67b50d870	17	0.0845266956187741	TRUE
8fcdcca36d7698798f6aa46b028ac8f7	9	0.0864096619651255	TRUE
b0d5a4ec8e74ac4c6cb4ecb9406183d8	9	0.0864096619651255	TRUE
415033681d70ab871d00f8e199d5485d	9	0.0864096619651255	TRUE
3dd272f8e982d693ab208a7f3d552074	9	0.0864096619651255	TRUE
536bca28d3d5384a4b1880b427e53dcc	10	0.105007314060969	TRUE
5c96ec70d2839b27eb48f2226c96e2ca	10	0.105007314060969	TRUE
4a765a31e0cf23936b0d98af5f0b7b81	3	0.115173403255331	TRUE
9cd7f253e5f398b99b6007964de56e802	3	0.115173403255331	TRUE
006447e914f38b29a86b92aab7b8145d	3	0.115173403255331	TRUE
e07ec8b4bb1731226c8bbe954aa9dfd7	3	0.115173403255331	TRUE
495baebe669e1a07a3638212d694929c	3	0.115173403255331	TRUE
b32b830756e43af68ac23a4f88d98118	3	0.115173403255331	TRUE
ba14674a604faae7a83e0ab9d7151dcd	49	0.115653659271681	TRUE
df6485f0c59f870ffa34796e56205884	4	0.153031417037874	TRUE
6836e6f80ca06e620471efbe1f47dffe	4	0.153031417037874	TRUE
2700db6dd1459de93b072b5f8f90c146	4	0.153031417037874	TRUE
8089fc5cea667447a21f0c665b788e9c	4	0.153031417037874	TRUE
872d11b3c9c9e72c129296b6ab82e76a	4	0.153031417037874	TRUE
776b3cb9d774b6e0c45c45ccc1764f1b	4	0.153031417037874	TRUE
0d87b112bc336b97248bf4757a668689	4	0.153031417037874	TRUE
1bfddc56f0549c7953359140bb8c64b1	13	0.166895948842549	TRUE
d372b6450fe1d69ae6cb673dfafdce6d	14	0.18905177036381	TRUE
490a5a0203d6a3ce512fe7566a89299c	5	0.190342995483322	TRUE
546a3202fd799a7240d2592ac0444502	5	0.190342995483322	TRUE
1b18df940027dc40b173be56b5b2c39a	5	0.190342995483322	TRUE
8cd7646d4eee2b702e680c9a36aa93ee	5	0.190342995483322	TRUE
e3064dc22ea1c1084cae7e065ec8305c	35	0.19916119737622	TRUE
074478577353df0840cde9827351d2e9	15	0.211740989485119	TRUE
a022ba393ceb3a7e109b171c5e901b31	15	0.211740989485119	TRUE
1189d4ef3af60ba943cc73161a4097a9	15	0.211740989485119	TRUE
7e06f6f4c9c89933a432f36767cf8f5c	6	0.226978141364239	TRUE
16357c104069b5c62302e913480dce46	6	0.226978141364239	TRUE
127b623fd88482c6bf125beda04dba52	6	0.226978141364239	TRUE
8b2b5fea59a9f7c4b75b285985f8bc45	6	0.226978141364239	TRUE
9cd232a91d762e72b8695ef807e4b595	6	0.226978141364239	TRUE
56d2f08c47e1fe5e8b11c758e3df45c4	6	0.226978141364239	TRUE
7cbcb853e8b0fc4e4ae9ad8a8893a7df	6	0.226978141364239	TRUE
42ae34a9e758d6ba2c23b9e32be90e80	6	0.226978141364239	TRUE
d0fd1a04107c82982b6555553b7bcf41	6	0.226978141364239	TRUE
b558365592f7a8d645f69873f6cda354	6	0.226978141364239	TRUE
e78a758dd0cf4ecdeec205371b241817	6	0.226978141364239	TRUE
262dc7bf78e937b0c7065aef82e18d08	6	0.226978141364239	TRUE
53e1b42a26f12c7056f1acb645630a26	6	0.226978141364239	TRUE
74d8b8501ebe056c7685c5f09a15ad1d	6	0.226978141364239	TRUE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

6e5eea14c8f71331e69a36f9df2e47cf	6	0.226978141364239	TRUE
e52a5d6f0e2560fe327b2f502dd5a024	6	0.226978141364239	TRUE
8aa02408859b45f0f98146372cadf479	6	0.226978141364239	TRUE
2cb9077e3906b6434cc06d7bf5c82e57	175	0.235878918358532	TRUE
6145f51f090994c4a43f7cea73f90091	17	0.2582325274065	TRUE
e269d8c81692a109f5be5e0d96cc45f4	17	0.2582325274065	TRUE
d5cd3fcafc36006692274229875ace16	7	0.262828389827809	TRUE
168860bd1024ae270960f35b81de4cc3	7	0.262828389827809	TRUE
029af632b2cc681c2baac604bc1d0f87	7	0.262828389827809	TRUE
7840fd72d5760bfb9d7c7d6e05e15b1d	7	0.262828389827809	TRUE
ca3d5ad542f41c1b2dca122a276dd68e	7	0.262828389827809	TRUE
ec5b70612ab541270f99ac9ccb03c57a	7	0.262828389827809	TRUE
5ec80473635e19defc7cb89911f32028	7	0.262828389827809	TRUE
bfe54af4c9180d37a0d76f6dafe79a5a	7	0.262828389827809	TRUE
527f393f41091fdc92723873b3f241f4	7	0.262828389827809	TRUE
4dfc1ccde14bb2c6f22c28071d19002d	7	0.262828389827809	TRUE
efb8602f3d0765aaf09b975799823da9	7	0.262828389827809	TRUE
5406ba4d58928df06e4f23d5a090f3d3	7	0.262828389827809	TRUE
65a39bc08ea93ac99b021f4dc1850237	7	0.262828389827809	TRUE
bd2ebc70501f7d867c204f94c4e483da	391	0.277100555069908	TRUE
340d2fc0f9822e35f88b16537341315c	8	0.297804452748937	TRUE
8c416f8840299049567cb009734c133d	8	0.297804452748937	TRUE
668fdb718997fc1589c7817655d4bb5f	8	0.297804452748937	TRUE
53d5533a6429e4aa7d6998450efe4a0b	8	0.297804452748937	TRUE
1ebadfa667c9a909e90f32ae471aee3b	8	0.297804452748937	TRUE
4fbbe364adcde3a84f51a161ebfa54b7	8	0.297804452748937	TRUE
7124e1c5c05a47eade99eb30e14ca10d	8	0.297804452748937	TRUE
b2487c386ddc1a950ba23034e0d108ba	53	0.298909268067355	TRUE
a71d241034f123d92d24606ffa097445	19	0.305485539889719	TRUE
d8ecd05fc246253fa7171e885a429464	20	0.329158787407426	TRUE
986d1160fa2c6d6e1b39adda581cbd09	9	0.331834080910981	TRUE
b474087ebf36a250a0d6e688aacc9b18	9	0.331834080910981	TRUE
de0880c59770b57d7bce7a58c29a63a4	9	0.331834080910981	TRUE
37bea876591b5db836c3cf500aeb2c4b	9	0.331834080910981	TRUE
743877d0dd17489090dfcc618262fe52	9	0.331834080910981	TRUE
2967fc5e7264a9da63556dc7aa011846	10	0.364860125823334	TRUE
29edd9d99bdeb69d7f7425b1fd0ba57e	10	0.364860125823334	TRUE
2e1a0c95b8b627148e4efc1a2e2124d3	10	0.364860125823334	TRUE
0bb1c8431343726b5826ba41bceb2deb	10	0.364860125823334	TRUE
69c82a0446a784e0f628baba67a71a38	76	0.379808917624895	TRUE
73bf8d1a5983e34a0cb84e3cae127815	433	0.382965201505468	TRUE
ef5b1d86471ba81aeaf0675cb5f133b6	11	0.396838784394777	TRUE
22c8e6e55e7d3f64fc2fdf7b5cc181a2	11	0.396838784394777	TRUE
6b0ff3b95f5a0170f8cf05f2d648be69	11	0.396838784394777	TRUE
0047d48fbaf4753f25019037ff802349	11	0.396838784394777	TRUE
c48ebcb64a879ac3131a8157726e0ce1	11	0.396838784394777	TRUE
5d6ee23084c6b9c96deb9a83295abc8a	78	0.408454653042303	TRUE
e370166a24f9ddb12cc92317e714f124	12	0.42773801098847	TRUE
44beba34c62eeb03966645efddedf0d9	12	0.42773801098847	TRUE
a32083174848f3d175ed977dc8eed9e9	12	0.42773801098847	TRUE
f5d0c756653e860f49eb2ef329840519	12	0.42773801098847	TRUE
c6357b5b5bb8c067c51faad5180b4e99	12	0.42773801098847	TRUE
40e44aaab0f468d14cc6f40ebef19cb0	12	0.42773801098847	TRUE
b1e1799fee04670207c1b71957c46d15	12	0.42773801098847	TRUE
03d52421a28d9b73c5d1716eb520718b	12	0.42773801098847	TRUE
22357830cad050ff79bf4e8f8488beb6	12	0.42773801098847	TRUE
147e062a32552d807a7a72da18b5c56b	12	0.42773801098847	TRUE
508524d40fa1145d6023b5a43a013f4c	12	0.42773801098847	TRUE
70e81e8f3ec3fdc237118834836a5ff0	12	0.42773801098847	TRUE
ab4429b988a4f6fc70f599b57a966c47	12	0.42773801098847	TRUE
677929eac52be081cdae08a0b9c70eb2	37	0.435706683699951	TRUE
55e6f47ff8b1dc2064c0182cb4f831a3	13	0.457536082595734	TRUE
b68071682d3bd6beb1a552c2500791cc	13	0.457536082595734	TRUE
48d9602985017b3a08622909440fccc9	13	0.457536082595734	TRUE
a36165e53c5432cedc59fba58c989704	13	0.457536082595734	TRUE
e3acd1847381805fd9cf7f67f238c142	13	0.457536082595734	TRUE
7d17d197d7978be0b6684209ce257e20	51	0.460922442542212	TRUE
1f4f4a82b05eeccd5361047345201e66	26	0.467192815197644	TRUE
f81aae39d658f14666bbd39fe410294e	26	0.467192815197644	TRUE
165cd1982e320b427a4cfa307f787b38	14	0.486220303988033	TRUE
f90e88c52df1bc147bfe9705270cb88f	14	0.486220303988033	TRUE
6f517a2da49d22439a74d5297bdb654f	14	0.486220303988033	TRUE
f342485eb28655371ab95467e436666b	27	0.489003842984167	TRUE
436e7f204759c2f369afb864438d1262	14	0.5	FALSE
c205d0ee4db8c2c892ebe949b0a069da	67	0.5	FALSE
38a149a8cc59529f0b4ce29d6a838176	162	0.500000000000001	FALSE
7fae0a4914d95fd2426aa8a2ea9ef487	236	0.500000000000001	FALSE
0cbf8d071fa1320c4a225b51196da078	85	0.506863732942227	FALSE
addf717222210dc7bcd173e43b0aae60	28	0.510360069573429	FALSE
9ceae8fed3a46d7b83b2131a74cedb29	15	0.513785840746033	FALSE
345b2e2f68d0022bf201b41dbd25743a	15	0.513785840746033	FALSE
7ed82ab99601f839d82875a9fed6caec	16	0.540234669026948	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

3081ffab8e3fff703f35599fbe5571c4	16	0.540234669026948	FALSE
ce0fe9dd4a105852b029c2ab1d997024	56	0.541776569825547	FALSE
c112b3dc575ef3b721464d0b49daa905	43	0.545307707097171	FALSE
0dc8e0e89e89cf4f0c71808666f6c6e7	30	0.551591656513902	FALSE
f03dffd2e63efb9a0dbc61f7ffa6fe0a	76	0.564343462092671	FALSE
44c386760dcb28e8af3b9847120c5d76	17	0.565574631822175	FALSE
1b158b8b2922d4fcad5d9cea607cbb7d	299	0.568562781856852	FALSE
ee32615285e3f0dfe911611dbd723a22	2	0.574036511156187	FALSE
bd8963179f96afda80d644b420b327ef	2	0.574036511156187	FALSE
495cda065a2ac9820db531b20e15aa1a	2	0.574036511156187	FALSE
59ebca98884af4ca79e99f56319810e6	2	0.574036511156187	FALSE
651b456fad22829867a61aff2a233caa	2	0.574036511156187	FALSE
40e9bb5ebb3c3102bb2bb4efe90a3a91	2	0.574036511156187	FALSE
2af0887c08fc8372e9894be7dd331020	2	0.574036511156187	FALSE
fc49e88a96a3582603047d56f9834715	2	0.574036511156187	FALSE
b76be88631fc1577ea6251a4ea6ad5be	2	0.574036511156187	FALSE
ece6618fc76d78380d6aa565960048f5	2	0.574036511156187	FALSE
746cc895942c2f16e6f9e05d9a912f00	2	0.574036511156187	FALSE
d3326adfc4e7e9d8ea929cc2b5ced9f9	2	0.574036511156187	FALSE
ea60a6f4d4e904f8f2d74612da65f602	2	0.574036511156187	FALSE
569333c671b24e2363d54e4671bb4cfd	2	0.574036511156187	FALSE
eadced2b44e8b968c557dfd8b5ac02dc	2	0.574036511156187	FALSE
329dade18db9c95dd219f5689ee777a3	2	0.574036511156187	FALSE
2613b77df8ce69292ad18bac50100eb6	2	0.574036511156187	FALSE
45c9cc46f257ba8de190eedb3dda5a36	2	0.574036511156187	FALSE
5c596a361996ea8d199802e4e34a6dee	2	0.574036511156187	FALSE
5bbe0a3078b4537152c396ef08fe6e78	2	0.574036511156187	FALSE
fbe5f855d23b47ec4e757e0fa3818ce5	2	0.574036511156187	FALSE
acd5713a39bbcb804b886f98694671b7	2	0.574036511156187	FALSE
3a247cedad2882a4fd1104f0ab9a65ca	2	0.574036511156187	FALSE
0bd637004516b97302a1f56e4f4d337a	2	0.574036511156187	FALSE
146bc18aaafdf9fc303f5cc2870c8e52	2	0.574036511156187	FALSE
bd4660e763464929bf79b4eeea699730	2	0.574036511156187	FALSE
c54410317ca2d9aa7aa88daa1147b46e	2	0.574036511156187	FALSE
dcc015b723865292b60a75224f821f71	2	0.574036511156187	FALSE
0360a55d85a1a3707023ff92dc73f805	2	0.574036511156187	FALSE
49b154d4c99e12186f73403e5ee94e83	2	0.574036511156187	FALSE
8c186c80797f8c256dc61aefe23d3411	2	0.574036511156187	FALSE
bd6a338d9c0a72f4b1a8410f16b85232	2	0.574036511156187	FALSE
a2068af49962d42e50ab5c8a0b744b71	2	0.574036511156187	FALSE
2f08fd7c00722cdaeb01512616c093a5	2	0.574036511156187	FALSE
b76404fb4d67434e693101ad1595767c	2	0.574036511156187	FALSE
e20f803f58f5ce46feabb0516cd02431	2	0.574036511156187	FALSE
acffff1db4fa874c7ff9e2825dd29a4b	2	0.574036511156187	FALSE
39e98dd4e807cd97e92b9f2169cafa9f	2	0.574036511156187	FALSE
791618ea7f777b9c2e7f908eebc1f2a1	2	0.574036511156187	FALSE
919de377b64f65dd94f18d0fca292bac	2	0.574036511156187	FALSE
dbf82288883244531ea0b57e2ed04a2a	2	0.574036511156187	FALSE
6b4fab5c304a37dd85146749d0028d87	2	0.574036511156187	FALSE
66c196e506f48ac22b611efac4a498d7	2	0.574036511156187	FALSE
981823f8c26086d499b6917a34f5961c	2	0.574036511156187	FALSE
20cfc8421a8c7570128653af9f9e5c63	2	0.574036511156187	FALSE
9d73ad67d5a5acbf511c78c859272821	2	0.574036511156187	FALSE
e5abe350b1c88e1126e9074c6e93ef15	2	0.574036511156187	FALSE
b3d6fcd212a9fd5fde39a3d79fee2faa	2	0.574036511156187	FALSE
1730d6157fef2851906040de8a99d96b	2	0.574036511156187	FALSE
dc40f19361b9081547655a601487f5ae	2	0.574036511156187	FALSE
3ed213ab80d34478df6c82075098049d	2	0.574036511156187	FALSE
ccea7446425ef2321bda5cb26650e0df	2	0.574036511156187	FALSE
97e91a44d9196a4f419d23186ef530bc	2	0.574036511156187	FALSE
480018a3dae196593b9153ec81e6d9dd	2	0.574036511156187	FALSE
e13ef43b8d2b3a52b013425e712bfb02	2	0.574036511156187	FALSE
a70bdbc2f74cab32c202b027329fef19	2	0.574036511156187	FALSE
bdb6ca07b4442f773d274bb49ef6a218	2	0.574036511156187	FALSE
000e38f477632fa6299901badf2e0f06	2	0.574036511156187	FALSE
70d3445bf16d2a93bcc75728b6d56c4b	2	0.574036511156187	FALSE
0c72ae13d05b1abb477b63055be82120	2	0.574036511156187	FALSE
c24ecd06ba3dbda7902bf9911bf372b9	2	0.574036511156187	FALSE
19981ee4c04815f6a74c02c8278b4dac	2	0.574036511156187	FALSE
6bf39213fe84babc512eabc37315923b	2	0.574036511156187	FALSE
473bc5c589d3abf37972dd9c7cde20bb	2	0.574036511156187	FALSE
34216a28c46c028e1026aae8d27271d6	2	0.574036511156187	FALSE
6f0deb7981650ae3e9ee2c309a4565e4	2	0.574036511156187	FALSE
963582325831c0cd60c85bb2a828935f	2	0.574036511156187	FALSE
dd14a2ab8cafbc7651bdb1cce5773535	2	0.574036511156187	FALSE
22d32c27c94d47482fd25cb0e8c65e38	2	0.574036511156187	FALSE
acc70fa94c2ae515818445e37268f59f	2	0.574036511156187	FALSE
35d552e9aa4e7126cc2489e0b0aab0db	2	0.574036511156187	FALSE
495174360f9b450d25be0fcdb035630b	2	0.574036511156187	FALSE
1a528eea8d2f33474bd7f647cc1ba694	2	0.574036511156187	FALSE
5c6ac5a27dd6c66c995e76f28ff3fea4	2	0.574036511156187	FALSE
9dabf344a87b6c26d226351cc10a2a5d	2	0.574036511156187	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

f59eb777bfe41f27f84f114609d558ce	2	0.574036511156187	FALSE
44cbe35f49cc2825a1cab0c45875238b	2	0.574036511156187	FALSE
66781e54d7c70751dd4315324ee6a730	2	0.574036511156187	FALSE
5bdaad19f2a370833f353f14c372764e	2	0.574036511156187	FALSE
eef731495f535ba4c0e34f7b6827b01d	2	0.574036511156187	FALSE
65a1f76dd20322568fa79a52f095c4d3	2	0.574036511156187	FALSE
7f0e8f6049e044128bb02c2904584fe9	2	0.574036511156187	FALSE
addb5446f7a31c8057158bfbe2e88a25	2	0.574036511156187	FALSE
c8f0d849875f193a6dfb78ce2f923fac	2	0.574036511156187	FALSE
d1a66fbb0ffb55d8ca5bde15afd2cbfd	2	0.574036511156187	FALSE
b75f8c1ee5be34e1e86f0794da0994a6	2	0.574036511156187	FALSE
d60b57fe3931d10aff6007227c714b94	2	0.574036511156187	FALSE
b524bdb0093f4363f6dcc957c3bc0cab	2	0.574036511156187	FALSE
4a32f1635f88214c4874eec85df56595	2	0.574036511156187	FALSE
1bc6c19320b6c5d08d6af72f39ca7028	2	0.574036511156187	FALSE
12c7139b414de8443898fa448413c98e	2	0.574036511156187	FALSE
59555b720479b4dae97998e8573280ce	2	0.574036511156187	FALSE
3f3a185ee907d56d95b84ccb84399854	2	0.574036511156187	FALSE
8c311c27fc6e8803ed957220d8645a52	2	0.574036511156187	FALSE
d8b253da31cae9272be8c76108df2c79	2	0.574036511156187	FALSE
11f72383f466cffd3d9c00fb0ddfa187	2	0.574036511156187	FALSE
6bf6bc06c4696e95747f2584057622c3	2	0.574036511156187	FALSE
5f0af45dcc3031ff73fa133f78f01dff	2	0.574036511156187	FALSE
d7059e9440d7d26e2e445d56303678d2	2	0.574036511156187	FALSE
59b13e3e404561803ca77c7fe3d03a0a	45	0.579476602130643	FALSE
69e611251f4d8582e312afa5737f033e	46	0.596027827052081	FALSE
ddbb3a2a7d6486a2201bb8770feb0a6d	3	0.606936130213229	FALSE
f5a382df02639db40c978149c852f631	3	0.606936130213229	FALSE
3c862616eb99d4336427f8c121410e30	3	0.606936130213229	FALSE
327af6c728df4ea81eb33ebb1c7b1f9c	3	0.606936130213229	FALSE
a256dcba109fe74b499e46900f7297bd	3	0.606936130213229	FALSE
8dc44508eb6a496df987c76cf6a23e10	3	0.606936130213229	FALSE
4ae0b8cee8d20168e08cb4f11ab7aeaa	3	0.606936130213229	FALSE
0d3677300a537a199d53c388dd4f9d14	3	0.606936130213229	FALSE
efea489e9aeea8ea8714c557cca5599d	3	0.606936130213229	FALSE
0cdc9b91d67a45a312f99f2e1a3b1462	3	0.606936130213229	FALSE
25817204502b01664ae7e0b787a23d35	3	0.606936130213229	FALSE
3d95c89ac99d998a8702babf9a83b566	3	0.606936130213229	FALSE
b6d5dae42d6441946b72536202970b6f	3	0.606936130213229	FALSE
78643cfd45f5d0e81fab3682a85bbe11	3	0.606936130213229	FALSE
6d51a149e8639a5edc3818184b07ef43	3	0.606936130213229	FALSE
cee08a446ff68f17526d7e203bbe871e	3	0.606936130213229	FALSE
e2dc6ea9ee0035962b6ecdbc882c648f	3	0.606936130213229	FALSE
ee496894344fde402c4a74ee69ad20fe	3	0.606936130213229	FALSE
ae31c634072e81af29cf4f857b5ceeba	3	0.606936130213229	FALSE
4675c263dc11b27a89c8b1a740a2b528	3	0.606936130213229	FALSE
445dc3e0e095d35c29018c23a15e9aed	3	0.606936130213229	FALSE
2605f4046444e0b60c1bbc901c97f065	3	0.606936130213229	FALSE
788883e3b7ee47924346c610b2ab6014	3	0.606936130213229	FALSE
85b4d7d480ef06ef2844b93115d87480	3	0.606936130213229	FALSE
8a0c8e17b96eaf823d7fc5b6ba2fa1a2	3	0.606936130213229	FALSE
3d1f0db0a5df73cc2642739c4ee2b98c	3	0.606936130213229	FALSE
3b80c7cd38deabc389b31026c8195e2a	3	0.606936130213229	FALSE
03f2666d9f52d8fccec13d113af7f470	3	0.606936130213229	FALSE
90eaad185a24ca78f31616c482b66874	3	0.606936130213229	FALSE
41997260ae3f1fa6e7c669acf6fa32f2	3	0.606936130213229	FALSE
8a5c6d9fa02edda0653c94cdf83899ca	3	0.606936130213229	FALSE
daefb5b04252122ce0f8c0b63169fd36	3	0.606936130213229	FALSE
7c57923c256db3ea4aa63f12e5afcc9f	3	0.606936130213229	FALSE
438dc2a2ea2405a782df5830ae62100e	3	0.606936130213229	FALSE
77dd072d2525de28b3d827121cdf463a	3	0.606936130213229	FALSE
e5816f291546ca05fda640ec24b5c7e3	3	0.606936130213229	FALSE
652b78d7b987b1efc1917822209ba690	3	0.606936130213229	FALSE
a99f6029c214a53baa6613443940b67a	3	0.606936130213229	FALSE
519d34f37ef0123b10f482f21b11bf54	3	0.606936130213229	FALSE
ade61be54713fafdfc05977a921ec339	3	0.606936130213229	FALSE
2fd73e273a5ab911a9bab5cec85e8178	3	0.606936130213229	FALSE
72ee964de9b0144e80347bd4e7112cf9	3	0.606936130213229	FALSE
1eeb8eeacd8184a41db687812348d06c	3	0.606936130213229	FALSE
a6fd733e44ef6fa8048ef0094de67f28	3	0.606936130213229	FALSE
cceb354db8986b34d8c5b435adfc4324	3	0.606936130213229	FALSE
9cab98d9be1a58a0486d67ec723285f0	3	0.606936130213229	FALSE
ca5e711152517b591d9d48b783f4942b	3	0.606936130213229	FALSE
ded5b76b4b3e821ef40d5df3fb48c91a	3	0.606936130213229	FALSE
65ff839c902a244022b20e77f4aed679	3	0.606936130213229	FALSE
3e5b4c5f9ec53fa58f406bb4872c2400	3	0.606936130213229	FALSE
579c3508fb7953140df95b7964b55fa0	3	0.606936130213229	FALSE
1289b5c9697f69da8b8f9648e728bc50	3	0.606936130213229	FALSE
0a9a722df3a13268be23ddee2cc2a0f6	3	0.606936130213229	FALSE
209ba71d1bcc4949b4d1fabd2ab1d5aa	3	0.606936130213229	FALSE
e474c7483a3a1001e73c7aeadf5b8ee0	3	0.606936130213229	FALSE
929d9a7f88bd83bc06d889c356d9a13e	3	0.606936130213229	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

47f1f3b66aafeaf2c4c7f9684ab7275d	3	0.606936130213229	FALSE
fb5bc4108d4ca6924d7bad409faff095	3	0.606936130213229	FALSE
94d028d1bdb01f40929f957b02afb576	3	0.606936130213229	FALSE
f889135f2c43218e4ddd55532ff76c3c	3	0.606936130213229	FALSE
85539c5100ea1d50723a53eb3099d883	3	0.606936130213229	FALSE
89b63341871a7ebd775e6f7064b1707d	3	0.606936130213229	FALSE
c1951d3e81e7f747aa634297072ea77d	3	0.606936130213229	FALSE
d71af0f02e2ca087c76c0a2b710c0945	3	0.606936130213229	FALSE
43152d438ac75bad2ec2ec1457fc6ac7	3	0.606936130213229	FALSE
0959782df74fea06fad1164f978fdd0e	3	0.606936130213229	FALSE
a3e2eb7398560e8e86867617d8011442	3	0.606936130213229	FALSE
a8195c108dec446f5c967078c25ebb74	3	0.606936130213229	FALSE
f79e9a78ab07909675c88e20cb0d95f5	3	0.606936130213229	FALSE
d92fcfcf215ad9fb193679b04a6737c0	3	0.606936130213229	FALSE
3136568cc32e1c6e2d0f114157863862	3	0.606936130213229	FALSE
7f68fd259107cfa9ac346ed09c5ef371	3	0.606936130213229	FALSE
13cdfd5b7da440810afbd5555701d91b	3	0.606936130213229	FALSE
3272f6b65b719378bb264a3740b2b40b	3	0.606936130213229	FALSE
fbb4b3994b2616a402100d1577ce4399	3	0.606936130213229	FALSE
55a225ab357fc87843d85d47fb8f2282	3	0.606936130213229	FALSE
c67d690916dc301d204b7ccb2bcc2ef8	3	0.606936130213229	FALSE
c1ae341aa77f14e7a1e3e62e413527da	3	0.606936130213229	FALSE
f39c1375b0565eda3b6660be85427c29	3	0.606936130213229	FALSE
7b6814cac868c1b645d0d8265e0fdd1c	3	0.606936130213229	FALSE
c6890b2e0e1c9ef31f93bff534d80700	3	0.606936130213229	FALSE
fe8d7e83f839d336881ec8419bb51293	3	0.606936130213229	FALSE
e0c17f6dde47055ed0caa4a7745e30c9	3	0.606936130213229	FALSE
02e0ce559228307ed79ccde3fe7b80ab	3	0.606936130213229	FALSE
151b90216c2f9ee35e184dafd9021130	3	0.606936130213229	FALSE
6babf4c5597da7990b82913ea7ac3ba2	3	0.606936130213229	FALSE
85a6f256aa6e1022a8e6bc648f01c037	3	0.606936130213229	FALSE
362d9d88f3195db02bbb580a020b0836	3	0.606936130213229	FALSE
b48900c23466c96cfaf4e4703c038566	3	0.606936130213229	FALSE
91090378e3a2b9e5e491fb0a627fb4c8	3	0.606936130213229	FALSE
72bb4b354b241a0bd9abd7caec6e9bae	3	0.606936130213229	FALSE
e548b7816111a16362f9dbfaf34be313	3	0.606936130213229	FALSE
eee80ff9268cc2e906d5b61786b66fbc	3	0.606936130213229	FALSE
b7dc1eedb15b4565b44c28c5fbbe49f5	3	0.606936130213229	FALSE
4e454938d9d1ae77d79ae8bf299a983f	3	0.606936130213229	FALSE
c56ac869661ff7a8d192fede12c318b2	3	0.606936130213229	FALSE
a4911447e5266f8a7d9bcef199267664	3	0.606936130213229	FALSE
1cf5017ed722b6a562933b84a7f7eeaf	3	0.606936130213229	FALSE
67496c58a44c58d60e084c6ae9f5ead7	3	0.606936130213229	FALSE
252e8532fee4f48af5e39f7514eb9d0b	3	0.606936130213229	FALSE
749b5b5cdf99a6ba907ad7e654702c7c	3	0.606936130213229	FALSE
b20f67d057bf23d44987c07876722fbf	3	0.606936130213229	FALSE
29a2ae2201cb196a28835102b6fad6f5	3	0.606936130213229	FALSE
5e94d08d6cbcbf35ab8d171568aa7196	3	0.606936130213229	FALSE
942bfb95d8f5f87f5a4340c64becbad0	3	0.606936130213229	FALSE
80565df1fb13bea7c5ffe8520f5add51	3	0.606936130213229	FALSE
9f7cd43f82859743a9fb6294b413138a	3	0.606936130213229	FALSE
6faaa11948882a569d8900983904d12e	3	0.606936130213229	FALSE
a4416bcf0ce3d2e25b3db4e987f4d88f	3	0.606936130213229	FALSE
7a1fbb0c1134ad5443dd743ea4b6f3ac	3	0.606936130213229	FALSE
5719b9b59cc1ec3e172cc9b9f3f0730f	47	0.612204436863422	FALSE
e59eb4f8c8550ddeb4ba43baf144ae11	19	0.612983675440831	FALSE
d56403223c8765b24812c19f1f844349	19	0.612983675440831	FALSE
6d3e98290cbed32480eb921f276470a5	19	0.612983675440831	FALSE
860c2d10b074762187a1f323f5f22558	19	0.612983675440831	FALSE
38bfbd3a2cf83da8fe03cea0f6b4f5f3	19	0.612983675440831	FALSE
760eb607a6f9eabf73fc3bc008113f73	19	0.612983675440831	FALSE
92ac8b57e0db3e849b7c8b0df7802639	19	0.612983675440831	FALSE
da1f560e36ff5523b6eae1acb26e7799	20	0.635090590400632	FALSE
49dc837f8fb48a7b37318967391f96d8	20	0.635090590400632	FALSE
834d07b2a02c48a75e52cd3d63fc7315	162	0.635262861075058	FALSE
9178843fe6c1b2b119e81a324c06e5ed	4	0.637356551907521	FALSE
79065bc5c081ea35aac41a59226dfaff	4	0.637356551907521	FALSE
857690bde71e50e3f9d91d2b77c8e675	4	0.637356551907521	FALSE
144e1bdc5d10e64c663d3833d416183b	4	0.637356551907521	FALSE
e40124c65379aab7ddc74c3d1c35213f	4	0.637356551907521	FALSE
6829b5946a95b929e9904067d572db6f	4	0.637356551907521	FALSE
7fd8ac343154ef6966a03f376782be92	4	0.637356551907521	FALSE
8419a2a1403b971c044d92d057e01ecb	4	0.637356551907521	FALSE
4313fef965a323b897e452786b227480	4	0.637356551907521	FALSE
4e7b984fa0360b99ffee1ceb10a122d3	4	0.637356551907521	FALSE
258f50eeaf72af21d70c648397c668b6	4	0.637356551907521	FALSE
36b69044091011464daa0325da37f61a	4	0.637356551907521	FALSE
dc5030eb18fe395a50b8235f2e583dbd	4	0.637356551907521	FALSE
4b95050efbe5bb490a63b0410c951326	4	0.637356551907521	FALSE
25e956ad84aba7ccf51f84544fe3b8d6	4	0.637356551907521	FALSE
3837a63a524a7a187121536fff409284	4	0.637356551907521	FALSE
1cd63754d2a45e4da3e2c17c07c39931	4	0.637356551907521	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

105e6198e75d958ec2beeb64369687e8	4	0.637356551907521	FALSE
42b8c9335747b0c86fc68d8adfce8315	4	0.637356551907521	FALSE
9654b46bd802f2c2748ecc5e77d97f21	4	0.637356551907521	FALSE
b7c98a016051abbbc602fa6163280ef8	4	0.637356551907521	FALSE
8c40ae61863ad1094a5fe3e7a1e56ee4	4	0.637356551907521	FALSE
e5fbcec9c3476c84ce40270b6f0bf4ed	4	0.637356551907521	FALSE
3cf8e3f2936c2215ad012e031a11dfc8	4	0.637356551907521	FALSE
5cc66b33870d22cf990e6a8ca618bb2a	4	0.637356551907521	FALSE
cad4e11393fe489023ae6a9724d6894d	4	0.637356551907521	FALSE
77f21f2da2d7b5fd7e7730b964e1ca40	4	0.637356551907521	FALSE
a509a05ddfc4a6df40cd9b1e75cb5f84	4	0.637356551907521	FALSE
618799965cb9fcfae385cc63346f300f	4	0.637356551907521	FALSE
bfd1c34f739028e507f085cdb991a0a2	4	0.637356551907521	FALSE
8d790eff268748c8792ab4fd07100e8a	4	0.637356551907521	FALSE
418d67cfd71b0f894bd6752f91a55124	4	0.637356551907521	FALSE
3f33b7ba47ff472f3dfa37699b3d4945	4	0.637356551907521	FALSE
fc1b76278975ee14760e108221ce14f0	4	0.637356551907521	FALSE
e56bd9b14d748dde0d00bfcba0c059b3	4	0.637356551907521	FALSE
ee5c297bd9fff3db3dca564f4de2b083	4	0.637356551907521	FALSE
392be618576de3abfe491746b24b64fc	4	0.637356551907521	FALSE
402e5913597695a16d7cad415ffff02f	4	0.637356551907521	FALSE
bfedbe08a11c47b4f1fba543a22b5d75	4	0.637356551907521	FALSE
0a9507c499d16037a527e88e32e7555a	4	0.637356551907521	FALSE
89623e037318e0f10a2b26123022dc25	4	0.637356551907521	FALSE
7f0a3c1eb166fa8b6b8c6169d5d2ee21	4	0.637356551907521	FALSE
5228eaef5a2de5ec5faecaf127b4488b	4	0.637356551907521	FALSE
2f1a6634dba0fcf6e04a10a0ab23bee9	4	0.637356551907521	FALSE
ccc5b420e1a3e5b891e8a467366084ee	4	0.637356551907521	FALSE
459a2ee9b40cae14ff8b1154046a45bf	4	0.637356551907521	FALSE
5ac0ca1b313546c59510815ec598f6d4	4	0.637356551907521	FALSE
9a1299533fdc202ea9dd1f8145e15c45	4	0.637356551907521	FALSE
1320e90134d8dd811304a92366e035b4	4	0.637356551907521	FALSE
7b25dfc72ef88a31085ef11f520cbc9e	4	0.637356551907521	FALSE
d6f114c9b544a83d9650ee4091d3c8e6	4	0.637356551907521	FALSE
6dcfd39fe967d9a976c330487e0ad0f0	4	0.637356551907521	FALSE
3cd349d8fddf16a5902fb77218065d47	4	0.637356551907521	FALSE
98976c388182f72eaf6444f252df2c4b	4	0.637356551907521	FALSE
f83d060cdb08f3a5c14531d80cb10d90	4	0.637356551907521	FALSE
7bb75a908f84cea53943cd1359b1deda	4	0.637356551907521	FALSE
1d4755131d0455f2ec5886e574308f11	4	0.637356551907521	FALSE
aba879ff52d1205f49a5c7c257679f61	4	0.637356551907521	FALSE
32e4f11674dad6bcd7909e3a195f3f3c	4	0.637356551907521	FALSE
f34bde4e8b265f13ee0a2e4b2e72532b	4	0.637356551907521	FALSE
39a7609929b77123cfe01080cd3a8ac8	4	0.637356551907521	FALSE
c6ff7bf11a21be39a971d86ac3f5ff02	4	0.637356551907521	FALSE
768d8ea38ea385a980b2218b8ccc6966	4	0.637356551907521	FALSE
391d3538ad8650017b13fe98bde557a8	4	0.637356551907521	FALSE
2178c86013f371f34ca8b6d82283dae9	4	0.637356551907521	FALSE
2bd44a9b4d580c5af16e7858b5554f05	4	0.637356551907521	FALSE
ece5f619b7415de3b8df0c94e1e708b0	4	0.637356551907521	FALSE
1b4572cc774d5173d5e09c03dbbca0eb	4	0.637356551907521	FALSE
88ab02c4a5b3b7fde63cdc6a21bdb6a5	4	0.637356551907521	FALSE
941e623512e7f598fc9942eaae0a6e70	4	0.637356551907521	FALSE
044e0092f4724300a52cce5c5fd4e882	4	0.637356551907521	FALSE
31921717848a5efcb875586d869f53f9	4	0.637356551907521	FALSE
71decf7564882efd1d2190fd29460d40	4	0.637356551907521	FALSE
16babc126b1836db96e5b5f9d59bc8a7	4	0.637356551907521	FALSE
31f7982d2607c5b3685d6c3974f2ed79	4	0.637356551907521	FALSE
8d4c51764c41d581fadd36ab1a7db29d	4	0.637356551907521	FALSE
30077f5042a2a53449f3afff0a7f21b7	4	0.637356551907521	FALSE
7cc6f23189123997fb56b704764660d8	4	0.637356551907521	FALSE
420001a74ca908712b4924c90f2fa112	4	0.637356551907521	FALSE
15a9309869945ab937c92fbe690903ef	4	0.637356551907521	FALSE
e503811c021343f5b11caed81d435f07	4	0.637356551907521	FALSE
4ce4bc11c99a28b0fe57853364b8451e	4	0.637356551907521	FALSE
adc1e763a2e0ce37c934f0ff2e25940d	4	0.637356551907521	FALSE
888f1e27e68fcfc06512e9bff356a16e	4	0.637356551907521	FALSE
b0d17476ed8f57285dc3fdf68d30d84c	4	0.637356551907521	FALSE
65e5847f2fe022a965dc8d4fca56434f	4	0.637356551907521	FALSE
06e10170152a7df3a946a557400c4bd7	4	0.637356551907521	FALSE
5e3736d7500f1847fde308557ffbcfdd	4	0.637356551907521	FALSE
fb3ace8db3915ddb8eecd36a0574f2f1	4	0.637356551907521	FALSE
63a00abba59283d87fa649e4b991da6e	4	0.637356551907521	FALSE
0a11bdf36506b66e44139ff447e2dead	4	0.637356551907521	FALSE
6c75cfbcda95ba7be4b96bd18203552b	4	0.637356551907521	FALSE
e0196d786487dae63ea300ec2eac2dad	4	0.637356551907521	FALSE
0ff344210a06866101d371ddca1193af	4	0.637356551907521	FALSE
41c984e77f65b6f4e3a141b213bb652e	4	0.637356551907521	FALSE
d51c2d5527534ea09b7c9279d8df7bb3	4	0.637356551907521	FALSE
8f62e1ce476fc35015f544e35bf6b791	4	0.637356551907521	FALSE
a10540eb078d6525325fbe1f2a607387	4	0.637356551907521	FALSE
a02a6468b6682c4ab0f0b31b8e080a30	4	0.637356551907521	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

341d0af40ca4d417d5d4b795ecb0cf34	4	0.637356551907521	FALSE
64301da69a74b94ac07fc109172f2c1a	4	0.637356551907521	FALSE
264e7ee6f0bf9bd1888f5b1bc40e18c9	4	0.637356551907521	FALSE
3e30a1f52cc81e130b8c023988acf9d5	4	0.637356551907521	FALSE
0ea6c0d1c3fc13d6a4d01653e44f19e4	4	0.637356551907521	FALSE
cf4b69428c4fbe0a55b81d9cfcf7604d	4	0.637356551907521	FALSE
3d06d03c32c7b6782a377199e96f576e	4	0.637356551907521	FALSE
a35640f17ff92ea37bea4104d9bb7f1e	4	0.637356551907521	FALSE
9fdb12014b64c13ab70cac6ab62b02ef	4	0.637356551907521	FALSE
803a31819695ae771de888afe8af7e71	4	0.637356551907521	FALSE
290d9dfd81df5f7192e22b2b347f4120	4	0.637356551907521	FALSE
1734c3e181110615bded5b21025d3976	4	0.637356551907521	FALSE
91f9cd9795f7b8793d9e06af80b0d56d	4	0.637356551907521	FALSE
0af9cf4f3e0a17840c19fda47ff11769	4	0.637356551907521	FALSE
202df81eb46cfff330eeeecc7e00092e	4	0.637356551907521	FALSE
ab61dbbc92d649f771918bab6ab7bbd1	4	0.637356551907521	FALSE
1056f4b6efc9797605ca25674fac9912	4	0.637356551907521	FALSE
45c63731d8dbb3b9e663139022627d43	4	0.637356551907521	FALSE
c1b4edf475fac73cbb145b05d22e6e3a	4	0.637356551907521	FALSE
b1d39025553d5a2e77eed0fd245eee5d	4	0.637356551907521	FALSE
0c3ca4bdbcc55a1eaf440eb73dd01a8a	4	0.637356551907521	FALSE
0a0f6a64b9bca840e46c7285b457c36b	4	0.637356551907521	FALSE
7cbbaefaca7a9bbb37376880e4a0f14e	4	0.637356551907521	FALSE
87118b7c468f54f714cd3a4701f728a4	4	0.637356551907521	FALSE
c6aae8ae72923155732f2e48a26b2fb9	4	0.637356551907521	FALSE
6359fb9ad7c1e779f94c583a2ec6fce5	4	0.637356551907521	FALSE
3bab94bbc9e095798d906d0afdc6d67f	4	0.637356551907521	FALSE
3752c0937de8a03bbbfc765819c78099	4	0.637356551907521	FALSE
c967baaf09da450bd29583233ea1a781	4	0.637356551907521	FALSE
c0c03235e8911295a37e19ba1c0af1a3	4	0.637356551907521	FALSE
b73ff1487c71f12a47318e8e9df59f04	4	0.637356551907521	FALSE
90422dff863388c4c097236cdfe729f4	4	0.637356551907521	FALSE
dbf3b32e6fcd7f6db830760df87aa435	4	0.637356551907521	FALSE
f2cc6eac9ecbcfa96ebd1c178e4ec6a2	4	0.637356551907521	FALSE
052bea98ea58d64bc7e0d792d4bb43c8	4	0.637356551907521	FALSE
b4e05e11ee1b2518be2326d1972121d7	4	0.637356551907521	FALSE
7dbb203fe0dfef7a33364e079154fc20	4	0.637356551907521	FALSE
d39d3f4f4d4881c0b75864dbe733de6d	4	0.637356551907521	FALSE
00f2f103259e0d2d279bec1899a683d5	4	0.637356551907521	FALSE
fc1c8932bfec8f0f632c63b5095bb577	4	0.637356551907521	FALSE
22872cce3904e3a7a73c15e4e63fcc2a	4	0.637356551907521	FALSE
de260172685c660a90923d290b078fd3	4	0.637356551907521	FALSE
38316233d8ebd5de6fdb0b87bdf0c285	4	0.637356551907521	FALSE
d482a8fc557bc969959b0c2a704c2c16	4	0.637356551907521	FALSE
79813aad3f28c522fd69a1e1e7f5892f	4	0.637356551907521	FALSE
ef628a53847821349bbe7988bcafb924	4	0.637356551907521	FALSE
85cfaa63370df8961eebc7b05ba076fb	4	0.637356551907521	FALSE
96416a69b1811ba8778fa127f20a3b15	4	0.637356551907521	FALSE
8efc96acc5fc0d9b6bda279be150150b	4	0.637356551907521	FALSE
e4ab5016c0eac0e4fbb1c8ea75b0ca28	4	0.637356551907521	FALSE
2aab0c408ab7e997e07d713b70cc5824	4	0.637356551907521	FALSE
6854a1092addd43549b415370cdb2d0f	4	0.637356551907521	FALSE
d31fb970ff9b6217462a4e0e52ca4b5e	4	0.637356551907521	FALSE
a88c5da1937f04d0c44ef7243e9476cb	4	0.637356551907521	FALSE
8cf45fe9e1251163b5d801841857a873	4	0.637356551907521	FALSE
a3699e0b9b431b8f11e1065820deaeb6	4	0.637356551907521	FALSE
901e24ec7d14fc177f7f9c15ab0f972d	4	0.637356551907521	FALSE
9a11856898d64d5906fca69779018c7d	4	0.637356551907521	FALSE
46e6034076b1f5aa371734d5c41605c8	4	0.637356551907521	FALSE
9232ae2e99a9314f625e80cc5f7a1b77	4	0.637356551907521	FALSE
15d8de3365192fca85622a7366050ee6	4	0.637356551907521	FALSE
8081227a3285bf886a9aeb207a890c38	4	0.637356551907521	FALSE
0cf93a42d2f2697d21739461b118851b	4	0.637356551907521	FALSE
3baabf8cc3e85c0f8b1d2790db20f838	229	0.647158144172996	FALSE
77d6a7db6de3bb1f15a02fdb36f90627	69	0.652990903333722	FALSE
eac36cc96b0e7277d4a485dbe6199beb	21	0.656163025642436	FALSE
d31af52f792634b80a89d3d310e619b4	21	0.656163025642436	FALSE
8f8539ce9266ce2c572b682f190cc810	50	0.658378182086215	FALSE
26c9f4f844da09bdd5c1603be193969c	36	0.662109791353658	FALSE
a1a3200b76bcd6000a0914892d370b6e	5	0.665479921351427	FALSE
442cc2db14651564914488bfdffd5071	5	0.665479921351427	FALSE
4c91e1ea3c1279e207a971df915d6178	5	0.665479921351427	FALSE
d6819c57d6aacc8113578b04b0911dc2	5	0.665479921351427	FALSE
79639ce66720319d718c17711fae3704	5	0.665479921351427	FALSE
8bfe73e7d3c17f599aec113950d751de	5	0.665479921351427	FALSE
501c1cf67a877c81057dd3fb330ddc18	5	0.665479921351427	FALSE
a412a5462ab0b34fa3b0b6dcdfa66b7f	5	0.665479921351427	FALSE
5ba826b6728c5f7a97293ac6c310a446	5	0.665479921351427	FALSE
e5ca098c7ee4d0fc9ff60e16e577c2b0	5	0.665479921351427	FALSE
03efc0161abb4e11e93f1870e9d5df40	5	0.665479921351427	FALSE
4b2862d8ddf6848328ccc64dbc6af0c2	5	0.665479921351427	FALSE
985e584faf3be7a41d1bca9d8e598817	5	0.665479921351427	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

56d75f22814ed2e9bb37c1452b26bc10	5	0.665479921351427	FALSE
37ec4bfd1c6ad9de0c6feef00d98d8a1	5	0.665479921351427	FALSE
800f4b5ec4543c7757ff29f6d5941d39	5	0.665479921351427	FALSE
ed40b08eedec4f873d38551aa43b0fde	5	0.665479921351427	FALSE
1dbd3332661515bcd3aef32a9b9e43d1	5	0.665479921351427	FALSE
621f0a4f8c2414b17f3b40434862a251	5	0.665479921351427	FALSE
c504e0ca94534572016a31761f44776b	5	0.665479921351427	FALSE
9b1a22ceee6ecfbdd684d46412f9dc49	5	0.665479921351427	FALSE
3bd445c971cbfbe4d6f12c07a68deec7	5	0.665479921351427	FALSE
72cef79e08aeb54df80915bab3ee3c01	5	0.665479921351427	FALSE
694ee281b908e4bfceb8d88584224367	5	0.665479921351427	FALSE
805599865882c8802a1ebbc3d539b68a	5	0.665479921351427	FALSE
bc4de31fa467e0b792b58e230e7cee7a	5	0.665479921351427	FALSE
51716d6182efdb29a158fe9562c83312	5	0.665479921351427	FALSE
71ad47c3b8cca9e31443d2370f72b32f	5	0.665479921351427	FALSE
d60a0925b76d5201afe16641d9cc1992	5	0.665479921351427	FALSE
2f267743f640a6790a563ddfbd2bf647	5	0.665479921351427	FALSE
d5c502312559578912b7c07f862f08a7	5	0.665479921351427	FALSE
d473f30bb25b09ce020fabb57f4ce14a	5	0.665479921351427	FALSE
a05f9f73ca792d7702b8dc6521d169b6	5	0.665479921351427	FALSE
99ffc98eaa2402b754b593ee3769ceb1	5	0.665479921351427	FALSE
382e9f4d997c89d86390ff1db5eac316	5	0.665479921351427	FALSE
1e5a5b3bc1ec5dcac873a685e4efed1d	5	0.665479921351427	FALSE
448092b5340d77b71bc09f3b5f6ba580	5	0.665479921351427	FALSE
90ce0670d0c5674699a2cd54638a8f13	5	0.665479921351427	FALSE
db6999c4df54a3eee3136c6671fa3517	5	0.665479921351427	FALSE
0b22ed48eb3acaf7d1d9721c8b83cbdf	5	0.665479921351427	FALSE
69a0a9b3d2f25e615792f49a88477b76	5	0.665479921351427	FALSE
e738239f47d0c43f0d59b1bbd3e5b262	5	0.665479921351427	FALSE
cb16531b22a650335c98dcdc66c2dafd	5	0.665479921351427	FALSE
84ec1399e8ee3618cc0686f28b89dbcb	5	0.665479921351427	FALSE
b5b5ccf8d4b597516a51069273afe7c8	5	0.665479921351427	FALSE
8ecb3c546a958d83e04d925a962af7f6	5	0.665479921351427	FALSE
9ef7688ddcbba5179a5011ec2c64a228	5	0.665479921351427	FALSE
08bb29e76019a1d1ab85c15f745c3b07	5	0.665479921351427	FALSE
f2c5e1386e7119686922dadc60904088	5	0.665479921351427	FALSE
6552f4787f1b6115bf9c9cf04a0b836d	5	0.665479921351427	FALSE
9cb31a8163c8b8dc83ddb589443f06e1	5	0.665479921351427	FALSE
96145f0f0d8025639c22921be63a5696	5	0.665479921351427	FALSE
be5ee53525a2da42755c32880c374493	5	0.665479921351427	FALSE
933096a8abba3ba26115619300c7b569	5	0.665479921351427	FALSE
13e78875b36fe62a172099127b50eca7	5	0.665479921351427	FALSE
213125af63c3c90e20e0e9e71cf103e5	5	0.665479921351427	FALSE
d32ab6db74a0b97252425a6b4d99b4a9	5	0.665479921351427	FALSE
3fa9b1137ccd47fc0b8a5960215cd91d	5	0.665479921351427	FALSE
b5ca209c65510a1622f68fa535441443	5	0.665479921351427	FALSE
e747680849d7059f0721a86f541a58aa	5	0.665479921351427	FALSE
fcdf3e1e3fd860c9a76a22fe6da59882	5	0.665479921351427	FALSE
7ef97632ab2d719134190b64d0e20481	5	0.665479921351427	FALSE
8c69b52248230f6661f7718a1dff9f26	5	0.665479921351427	FALSE
5c2fe78d3aa8a5952c3261269b857cc1	5	0.665479921351427	FALSE
3bdf3962a4d6b9738dc1bfa095f6cc8e	5	0.665479921351427	FALSE
d0099ee91dbd9830ee9ca81d0ee85747	5	0.665479921351427	FALSE
93da342980ee5d7587028b142860e326	5	0.665479921351427	FALSE
f8bfa38966ea4a31ede4d309eff80d6f	5	0.665479921351427	FALSE
1a7a8f466b74b6b5e3497d10112cda38	5	0.665479921351427	FALSE
0e86e0627c7da65f03262253a0f36a37	5	0.665479921351427	FALSE
5f8fcfd5eae442ef695287a85c782c5d	5	0.665479921351427	FALSE
b3f501eb848286a3758e9ce88eb51d01	5	0.665479921351427	FALSE
c9afd558a3687a3ab1cb10e0979c6e26	5	0.665479921351427	FALSE
f795c30a2d4c83ad90760a26b8948100	5	0.665479921351427	FALSE
763da75108822f3211c6f8852fecb09d	5	0.665479921351427	FALSE
0ea5e1ee42a9add6e20e4cea1e13e61b	5	0.665479921351427	FALSE
fc16f182cda13158b7b53b99ec439473	5	0.665479921351427	FALSE
937130634699adf72e663b3206296108	5	0.665479921351427	FALSE
684661c135c9eac8d14167f4c0f04940	5	0.665479921351427	FALSE
a2a66718d4e7cf14e217a76f6d99900d	5	0.665479921351427	FALSE
25b00fbc430d7f939bd949e6460a84dd	5	0.665479921351427	FALSE
84a73b80b8405e0ddf23fcd9f218acbc	5	0.665479921351427	FALSE
e5ba1449f96ae19c21538f0ebf01b59e	5	0.665479921351427	FALSE
2deee102fe6514204af2bf0780bc1ac4	5	0.665479921351427	FALSE
b987b751d2357f99e82610cc984b09e9	5	0.665479921351427	FALSE
5e6725e17a517802fdf48cba3bc05c1c	5	0.665479921351427	FALSE
2fc9b7df35cf9b721acedcd9d3561cce	5	0.665479921351427	FALSE
23a7b18fbefa1fa371830e7db7876b29	5	0.665479921351427	FALSE
62a87d4db851644446615062d61d09e7	5	0.665479921351427	FALSE
99d5a4ffa07fc3904d57784b335f19bd	5	0.665479921351427	FALSE
f0b16686cde37268349599e11a1f8790	5	0.665479921351427	FALSE
934f56716628cf3758baa5e7b702d2a1	5	0.665479921351427	FALSE
05583e0e12acb4607a40b2f17ce39e18	5	0.665479921351427	FALSE
aec20e1a147e3f84f9794c1b9066aae4	5	0.665479921351427	FALSE
3b09d01fdbc5a425b7e984a53f59a9d8	5	0.665479921351427	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

4581c03a9fbf22ca8d26d197945c5d06	5	0.665479921351427	FALSE
0e089dac21915d34e1ba6011443df2e9	5	0.665479921351427	FALSE
09b47236ecdcbe4627b4f19dadfb2867	5	0.665479921351427	FALSE
57706cd140992bbd0d4e96a833912f45	5	0.665479921351427	FALSE
0d2a4db996f43b65f6249c118c8d424c	5	0.665479921351427	FALSE
1dee7a9043af345aeb0c01ef1f4473ae	5	0.665479921351427	FALSE
0f21e6a5615139dabd413632d1378dc6	5	0.665479921351427	FALSE
9bf47859277b18f7164e848576d8829f	5	0.665479921351427	FALSE
9243a3b81d2e0ffe7466f757feccb47f	5	0.665479921351427	FALSE
687bb3b3b1b09601bd3a0ac6eecaf46f	5	0.665479921351427	FALSE
dff5823839d97156151935cda410bb6f	5	0.665479921351427	FALSE
d6eb49be064bb019a8dbb28451f649c3	5	0.665479921351427	FALSE
0bab6eefa020ad40122cff733972d7f3	5	0.665479921351427	FALSE
7b732016b53fc4c4128273ac1c83fae2	5	0.665479921351427	FALSE
6875c5259a37bc8844c092db19fec963	5	0.665479921351427	FALSE
1a2d11c18878fcdf1192fb22e7e45a93	5	0.665479921351427	FALSE
637acbfe700723da791ff81b54634740	5	0.665479921351427	FALSE
4fe609069ddb659da47531c688130164	5	0.665479921351427	FALSE
7b897f5ab602ed700ec9c10dd96fb997	5	0.665479921351427	FALSE
1b097ec2f993f8d93f006bdd6499a2b9	5	0.665479921351427	FALSE
3615fb777005e792f48281d9766a159a	5	0.665479921351427	FALSE
0b49776f60bc0f566ebe8585ec3c5e07	5	0.665479921351427	FALSE
c45fba966a7c86503b38a13e121fad0a	5	0.665479921351427	FALSE
e39d0106129d2cef95d226ab68163276	5	0.665479921351427	FALSE
69e91084b18e0bcb7404b6a55fc3b17f	5	0.665479921351427	FALSE
e30af441c32096bbff179e10a8d80c75	5	0.665479921351427	FALSE
a3fdee62edf9cb50662e84281124a241	5	0.665479921351427	FALSE
12b3baee9905c9965794265d33f8511e	5	0.665479921351427	FALSE
306a0dc859c17f4586aac1cbef4f7d45	5	0.665479921351427	FALSE
af7584c5080425da5c5b2bf7a2045749	5	0.665479921351427	FALSE
8fea6fb414f3973c60bb3d9c9b396391	5	0.665479921351427	FALSE
858e61f4b80d4c21c57fb23e4adf00b9	5	0.665479921351427	FALSE
f42fa1f361123808c5170776beea4f28	5	0.665479921351427	FALSE
a3d55180faf1984a6ecd4e4cf7cd9621	5	0.665479921351427	FALSE
833e68d8409bfed100f147d62868285d	5	0.665479921351427	FALSE
bada4b4d5ec4bb02547d109017bddbf6	5	0.665479921351427	FALSE
1de11ca9bfc49d2a24f3217092474de5	22	0.676227110786361	FALSE
e5d49fa90328817f763cd8e8f5b0eab1	22	0.676227110786361	FALSE
6da46d83d8eca17871b9937d084ea890	22	0.676227110786361	FALSE
bab727aead750abea99848d1f7b760d7	71	0.678242374515425	FALSE
3b482b3f9da81db80a221a66031aa159	113	0.684135587722467	FALSE
476738e317c846bab51ab3cbefc92f1d	6	0.691475346686081	FALSE
39ca85a8156900264d039677c05ca2e8	6	0.691475346686081	FALSE
db0a9bba4758c990e4d72f8558596479	6	0.691475346686081	FALSE
c00181a21a8c5b8b22d1b2793f6fdfb3	6	0.691475346686081	FALSE
0984f915f803bc5c24bf8e1c909ab5ca	6	0.691475346686081	FALSE
81d51386cb12c21a2f27ebdd8b9b2ccc	6	0.691475346686081	FALSE
375f634194a84d8e75ae82cf5a65769f	6	0.691475346686081	FALSE
642cd6c8b3c6a14c90307513a1249ea7	6	0.691475346686081	FALSE
ff4647415dccadee59cb477abc5de829	6	0.691475346686081	FALSE
771eef8a8acb8fe4bab03d049dd070fa	6	0.691475346686081	FALSE
fc48628625c6ec016a4d57e4e6d3d5af	6	0.691475346686081	FALSE
a810d316616fb1a7321351393c70ae70	6	0.691475346686081	FALSE
42adb322e812d7d110c13bc5c582464f	6	0.691475346686081	FALSE
f54538c9fe78a2491a7365d46ac7d04a	6	0.691475346686081	FALSE
5440bb497433d6fb4d592fc9dc4115fc	6	0.691475346686081	FALSE
846d9bbb062cdef518cf911cb5f2f48e	6	0.691475346686081	FALSE
bfd5a200f7cbc8ed661fb5e5659b809e	6	0.691475346686081	FALSE
d7d8e9ca0bb47e48bb7b11010bf44cd9	6	0.691475346686081	FALSE
19032423b2abbba6747425ffe54db960	6	0.691475346686081	FALSE
24578f723e54aac8e9d917365bcf8d87	6	0.691475346686081	FALSE
cf82d65756dce349e1dfeb64ab983821	6	0.691475346686081	FALSE
9c6f2b50ecee1d000f4981af2d9ab517	6	0.691475346686081	FALSE
b7a90c7afe20bf32a35c949e18984553	6	0.691475346686081	FALSE
2cefc396eb5f251d5f5142811640f757	6	0.691475346686081	FALSE
128494a261f2bae571bb1058b52c687a	6	0.691475346686081	FALSE
b493b2e590a925aa8aa5e8e5ef559dfd	6	0.691475346686081	FALSE
6645ac6e01b54dbf89dcb1a3515e387f	6	0.691475346686081	FALSE
eeb6eda609da001bc1cda61fcdb1f67b	6	0.691475346686081	FALSE
0fc819b99aeeaa0affa6d9db40378b17	6	0.691475346686081	FALSE
d89b8bf4cf0e4e17fe46c0d214238584	6	0.691475346686081	FALSE
20ed485a649cc0f03dd3cf2b9a488a22	6	0.691475346686081	FALSE
74df131c4c31e7469f7f71f4d93cf57d	6	0.691475346686081	FALSE
64a76d1d022bade634287ace4305d5c4	6	0.691475346686081	FALSE
99b0ace8f5564cc9f51d62e073d1389f	6	0.691475346686081	FALSE
c05eef21fa88381de2bcb04b003bcad6	6	0.691475346686081	FALSE
38e37205b8f101040fafa39445a00687	6	0.691475346686081	FALSE
7e28ca07fa82d553ff3e2eb951f96b1e	6	0.691475346686081	FALSE
2862752121447c336cdaee1089c3bf9d	6	0.691475346686081	FALSE
cfeaa32bd4779c561a79d9ca355a8a21	6	0.691475346686081	FALSE
d6a07a207695e918c576157088e00158	6	0.691475346686081	FALSE
69d658abc0bc72ff78e012446a0620d0	6	0.691475346686081	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

78c77fe65ef85f77876913afafaf83ca	6	0.691475346686081	FALSE
21f57b10a24af2d07ef8bb29fde77308	6	0.691475346686081	FALSE
fbe1441247a9f5ff2154e296ce755b02	6	0.691475346686081	FALSE
19faab0613301c627436bea7ba130c03	6	0.691475346686081	FALSE
c372b43be47a60debddb8ba64f5c4f3b	6	0.691475346686081	FALSE
4281f5b1a9d4da51b4f84a86734e3016	6	0.691475346686081	FALSE
33fb7c931529eafff0ba2653fe5e55a4	6	0.691475346686081	FALSE
54eaa3dad1c82a5c1a864cb36dfa00a8	6	0.691475346686081	FALSE
09654246eee58793bd55f4f2e8866c94	6	0.691475346686081	FALSE
1554445ec0f58f8f3aa89dd4c19f6ca5	6	0.691475346686081	FALSE
725c3ce493fc20f290177b3cd00ac157	6	0.691475346686081	FALSE
d17be2e7da52698326ae236a7e80bf0c	6	0.691475346686081	FALSE
5f8ae198552bc39633ea5c3962946313	6	0.691475346686081	FALSE
07b2282aba42b7a552743011db8094f6	6	0.691475346686081	FALSE
3e9db007ba1ad11c10840b072cf82ac9	6	0.691475346686081	FALSE
ba2abd56e0c62e84a281b252c0ecf2c4	6	0.691475346686081	FALSE
ac58e247b4b4dc9952a63a8768842531	6	0.691475346686081	FALSE
f0c33566b57c73834e6715e4f843280f	6	0.691475346686081	FALSE
3447bba88a77090a6e1f60be2df80a8d	6	0.691475346686081	FALSE
c9495d05babd9ff3d12861c52ac48b10	6	0.691475346686081	FALSE
cf1c154dda5617a313977704c70b4fea	6	0.691475346686081	FALSE
665c06c0446675030ffb92788fbc2972	6	0.691475346686081	FALSE
de755d7719f17159286b670fb8ec8452	6	0.691475346686081	FALSE
f4009a122f9d6677ef2cc595acd70db1	6	0.691475346686081	FALSE
023f6419e31912178c881067c038cbe8	6	0.691475346686081	FALSE
0ba7ab54477ac1e358b95b12ae47cfcd	6	0.691475346686081	FALSE
e5efb19c8e76cad58107d2acd1913b91	6	0.691475346686081	FALSE
282cf87787bf8bc13d8469ce93ea9a83	6	0.691475346686081	FALSE
37593875f248af0a9f8a410adae37818	6	0.691475346686081	FALSE
e53ed171ef0054d5828f13109e670473	6	0.691475346686081	FALSE
0122b585024361ddaee2c8d00f210974	6	0.691475346686081	FALSE
88a665d3728e9f8cb9660a1ce9b3c3bf	6	0.691475346686081	FALSE
05e59e6acfa67481514206aea60b5817	6	0.691475346686081	FALSE
ce83fcf423e41adb5b0ed8d2ef7c692b	6	0.691475346686081	FALSE
9c23f56aa5a85f3d6ff72ab45dd6a8eb	6	0.691475346686081	FALSE
e6bbc8eaac8196675f99a78761aefab9	6	0.691475346686081	FALSE
6bc1b948bdac137349310bd8e6822f34	6	0.691475346686081	FALSE
e41b96f0ca69f99a45f8985d840fde43	6	0.691475346686081	FALSE
4202b2ba3e89e89d60d5e1657251a277	6	0.691475346686081	FALSE
3a75c7fcc9371c1129d5e490d67b5473	6	0.691475346686081	FALSE
302e745422686c8b40fecf517ec856ca	6	0.691475346686081	FALSE
d1cb4ea8ba58a5e6d1be97779e72dd3b	6	0.691475346686081	FALSE
5b115c09c95c3a6b956d53d2f1884614	6	0.691475346686081	FALSE
fac0a19f742f17a596e1554f6f94542e	6	0.691475346686081	FALSE
92ea3b676285053413ddbf4880f2a4d5	6	0.691475346686081	FALSE
8f339184fac227ee4142a3554f5cc64f	6	0.691475346686081	FALSE
f2b214c47f5b29b3edb2894e3fd95dca	6	0.691475346686081	FALSE
1e85713d0ee4b69c1ba7b57fecce1e7e	6	0.691475346686081	FALSE
1ae8041320c459acbb58d90de8a1b346	6	0.691475346686081	FALSE
f44603f6b65d6c112a977783a13461f5	6	0.691475346686081	FALSE
8de8c318baec968219319f50f648af60	6	0.691475346686081	FALSE
67e330de0289f13cf25c1f4ac71dbf19	6	0.691475346686081	FALSE
9729d58a6a7dd9ddc2d0977cc7bb68cf	6	0.691475346686081	FALSE
2bbafcda6fea1b9cbbc8610409b06bf8	6	0.691475346686081	FALSE
64485352a43d56f50f1b3a5f3167b0a4	6	0.691475346686081	FALSE
1509e5bbabd9d1c71b407a0ac218e46b	6	0.691475346686081	FALSE
a5dcc6de4c99de98137e828069413f13	6	0.691475346686081	FALSE
04e24a19d69d50187392f3b293672907	6	0.691475346686081	FALSE
9da052e6c59fcdfcf34e1afb8beb933f	6	0.691475346686081	FALSE
abbdde2546c73e95a40fa010995d2818	6	0.691475346686081	FALSE
29cfe68fe8b380a7be28a396d32cb6ab	6	0.691475346686081	FALSE
725fdb050bbeefec541866c137e9c50a	6	0.691475346686081	FALSE
eb2594d058017aaba416b552608fa95f	6	0.691475346686081	FALSE
8115bfdc03a0ef73ed285dc9716fc0f5	6	0.691475346686081	FALSE
b62deda5ed36b8c96b0acc4011f0001a	6	0.691475346686081	FALSE
9f6e744b36e913bb4649329926fbfee2	6	0.691475346686081	FALSE
96714a6842a401aa4fecabff7a334c5f	6	0.691475346686081	FALSE
2a7d7d28b641f4285a458ca5cfacd950	6	0.691475346686081	FALSE
6cc51fe5dd109eb63633c531f2ffb660	6	0.691475346686081	FALSE
3a228aeb41029dad281ee5d1ad796636	6	0.691475346686081	FALSE
511bf33ffff1bc6d33e8b7024543008d	6	0.691475346686081	FALSE
8839f92e94f7eaa6b15070cdd3a35932	6	0.691475346686081	FALSE
d04db31eb1abae895d806b53ef3e6d33	6	0.691475346686081	FALSE
f018335ed27f4ebb1191e4b07133ba72	6	0.691475346686081	FALSE
8210bded00785df70ea9736d5be8363a	38	0.694321371129945	FALSE
e73d7f3f0353d7c592d2e13108551f31	23	0.69531093863704	FALSE
d881f447ded933c782b33f0ed029eb81	24	0.713444141883855	FALSE
6df3a5390512a230ba46b8f69c083579	24	0.713444141883855	FALSE
78bade5ebbe4174d8be9924c6185e63f	7	0.715499807394952	FALSE
000653cdb81bc81c8296b722a5cfaa0a	7	0.715499807394952	FALSE
849d52ff327105ded96e282fe1bd3b62	7	0.715499807394952	FALSE
35ff75125aac19c43b821d316def91a6	7	0.715499807394952	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

f6abd5e319f04daeaf391a2e5654004f	7	0.715499807394952	FALSE
30dbe645e8a27035b82f3e2724544d5e	7	0.715499807394952	FALSE
6e5b0b6f7e2bb31c463b644e85bde992	7	0.715499807394952	FALSE
6861dd72bd088689e5dfef66bf7e4135	7	0.715499807394952	FALSE
9c2ebf8f2921e2d7bb2e2270607342fa	7	0.715499807394952	FALSE
673173eec4db0acd7101cf10a16b6866	7	0.715499807394952	FALSE
18786531e995f35bfafe6938d96ec702	7	0.715499807394952	FALSE
af6374578fd7c52c9e70103c558b23ad	7	0.715499807394952	FALSE
5606b84bea115e968c56e015ba80be07	7	0.715499807394952	FALSE
6047e1cd01fa98962f888d9f00cbfa6e	7	0.715499807394952	FALSE
d1de072c491257d797a532babbf5764a	7	0.715499807394952	FALSE
15329d5033e2beb83d980271241d8308	7	0.715499807394952	FALSE
d1cfb51391363f63a2a4dae00098c508	7	0.715499807394952	FALSE
9f7c2bd6ef46996c8c194978d581c879	7	0.715499807394952	FALSE
5ce34ff294c4eb73dddb425696c47c60	7	0.715499807394952	FALSE
4cad7317878eea3b34ff87a0aae89c58	7	0.715499807394952	FALSE
75dcf4386d37d4fc0f8d2e6b17ecab8a	7	0.715499807394952	FALSE
1579f92d06c86388b8afd5ff14a3b211	7	0.715499807394952	FALSE
d7f26dea2af5078b6b71563ab7cd6eaa	7	0.715499807394952	FALSE
f9f6d2c29406a1846e6da618176b1e21	7	0.715499807394952	FALSE
cd46e01b7ade82713425c980e460e5ad	7	0.715499807394952	FALSE
426ca56efb6816b00bdf050ab01d0d0f	7	0.715499807394952	FALSE
24216962fb504c3ce0264bf24d1f9cb5	7	0.715499807394952	FALSE
84d7876514c79823177f54c6bad6c2b9	7	0.715499807394952	FALSE
bb7baf63968d36cf1bfbd5fcd580ed70	7	0.715499807394952	FALSE
8b85eaaafd6c5e4b10a90a4a32790e4b	7	0.715499807394952	FALSE
5d9b474ea44d8ca1f55007f87e53a13c	7	0.715499807394952	FALSE
a46eeca1b1204fe087c1f1f5b2e7cf55	7	0.715499807394952	FALSE
9b2e53b19d862d6826fbebe3d423e618	7	0.715499807394952	FALSE
8181046c6027b76fecba11accf9efe8c	7	0.715499807394952	FALSE
e95ef791d06b1d2f990345ee53ca21b8	7	0.715499807394952	FALSE
d3a26f6c315d1c86b14b88e81af8e3fc	7	0.715499807394952	FALSE
5c6edb849229155a7f835f12505003fd	7	0.715499807394952	FALSE
93b6acf12ab28286f735e299fdcf7e3e	7	0.715499807394952	FALSE
985c885e396f1e7ea883ebe31b3e3e66	7	0.715499807394952	FALSE
09595e80206d35b965f4cce02b2f48c3	7	0.715499807394952	FALSE
30658a6c43516aafecc4b553b81e6b8d	7	0.715499807394952	FALSE
4ec764dc072f2f639a74752149d1b9ef	7	0.715499807394952	FALSE
b6c6e727b8a2b38f3fe4f236aab51182	7	0.715499807394952	FALSE
d2851be2899896813406335f097e0766	7	0.715499807394952	FALSE
3b8a7748b36653caf4a339c462aeace7	7	0.715499807394952	FALSE
c219cba5ccd765a396f14d20423f4f5d	7	0.715499807394952	FALSE
b3c39c59af5fea7457921f190fddf246	7	0.715499807394952	FALSE
fddb21906df27ed56154006167998223	7	0.715499807394952	FALSE
2060cd013f6f8e1fdfda2e4fb5f9105c	7	0.715499807394952	FALSE
9068298bbfe44765abf5d4b36326b8f3	7	0.715499807394952	FALSE
f07c42e5c7ba021de7b60091ed440e76	7	0.715499807394952	FALSE
09c3e7640cd22403cb1b069579db6645	7	0.715499807394952	FALSE
cd615ef0610acbcbf4bafa1586ad08ee	7	0.715499807394952	FALSE
e0287daf63905b8cf23ad925f12009b2	7	0.715499807394952	FALSE
6dd8a4fb3215ac145571b0079405a312	7	0.715499807394952	FALSE
a116c09afca2ad940b8906ca4c225822	7	0.715499807394952	FALSE
8d8bd5de425b804441258f27801cd9f6	7	0.715499807394952	FALSE
75c37c29a6bc563618b656d63a21976b	7	0.715499807394952	FALSE
f40ca23c1724abd681edb4205954c756	7	0.715499807394952	FALSE
a8b203461e5cfead0462ac1dcfee07c0	7	0.715499807394952	FALSE
1c132342410d69d37d82a2aed3bfbb05	7	0.715499807394952	FALSE
dc41358fd306fc1053d80935c0ddf0b7	7	0.715499807394952	FALSE
073ce754d4fdcb4e4828e63e4a6387fe	7	0.715499807394952	FALSE
9bfc9ecb4864e67cd7d432a627d7fe11	7	0.715499807394952	FALSE
a84ac8fc73810467a41d4b54da068bf0	7	0.715499807394952	FALSE
169fb216a12f673fcdd8cdcb5de35a82	7	0.715499807394952	FALSE
00d9ad4346446bcd2b34269f283f39ff	7	0.715499807394952	FALSE
ee52da5d4bb76047ccf285398cad8f73	7	0.715499807394952	FALSE
978ba7cb1df2117169afdc062d0dbe46	7	0.715499807394952	FALSE
71c9b7fd6b750486bdf5c59543b98a6e	7	0.715499807394952	FALSE
f4d5b9520fb5e0c48fd3afb1fc9a5047	7	0.715499807394952	FALSE
c65b7d39ebce1feb1d58e27aee1bca47	7	0.715499807394952	FALSE
9c439916915fd3dba65264406eb263ba	7	0.715499807394952	FALSE
a91695ef6eb933fc1a78804ad848ca05	7	0.715499807394952	FALSE
5c046402b9d4eda1925b61e367712fbe	7	0.715499807394952	FALSE
a4591400fc7d7ddf8843573260636cca	7	0.715499807394952	FALSE
1aa75dad1dee675701b42a805fb0d42b	7	0.715499807394952	FALSE
a0999abb73d84e91d5134d80528ef229	7	0.715499807394952	FALSE
4d4c903ae6df84e7f79fbe3cecf19268	7	0.715499807394952	FALSE
d0b698c7298bf04110a6d2f220879bfb	431	0.725498775526386	FALSE
be895bf9176954ab8f55e2c2e1d4e047	25	0.730657519292985	FALSE
2ed220cab9db0862029af74d35ffbaf4	25	0.730657519292985	FALSE
d8cb744c811fa780dd595cf6df39bc60	25	0.730657519292985	FALSE
1357277b4f47cd8732a2707eca12fcdf	8	0.737699001068447	FALSE
170d6813dfbe0ce95dc78d234a5b7155	8	0.737699001068447	FALSE
bd6e442e02ccd1c8d74c8632a93d3648	8	0.737699001068447	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

518b140a9a79e87da975636f1e42f403	8	0.737699001068447	FALSE
e0de3364d91c93c955f117c62e66a9d6	8	0.737699001068447	FALSE
769de0c67929c3c1a4007c54642dfe34	8	0.737699001068447	FALSE
b5e089a90fc5455638f85de23bb20993	8	0.737699001068447	FALSE
80098ec13d5319e6024df222a32d16bd	8	0.737699001068447	FALSE
49c5631b22e3c8af1bcf6caf37d3ff7d	8	0.737699001068447	FALSE
62169a1df5dd845f1a7de084c2219327	8	0.737699001068447	FALSE
42322165cfa8b7b298ef6e3516d0f4a3	8	0.737699001068447	FALSE
97a36b566a19741b7e7ce108e99a678a	8	0.737699001068447	FALSE
11b24486ce2911727bd082a6d0bd868b	8	0.737699001068447	FALSE
6d8da19cf86ecb2c26f21b692f794071	8	0.737699001068447	FALSE
0ca19cdac2fb9198a038aff1281a5e10	8	0.737699001068447	FALSE
18a5046a2252d72a8f9a425127ee63c4	8	0.737699001068447	FALSE
b161a9890947c468aade8a7f27ccc163	8	0.737699001068447	FALSE
a8b23d869d9990033ef96272f0fcf119	8	0.737699001068447	FALSE
6de665187bcc484ed4efad4465646a05	8	0.737699001068447	FALSE
69eb16a95fa594304be5c161c6b15c9e	8	0.737699001068447	FALSE
f14a4043d5ce3e764003e5fce86495e0	8	0.737699001068447	FALSE
daaabb96173b74ef79fd1aef40e9457b	8	0.737699001068447	FALSE
3016b2fab82ed08bfb5585b175919df6	8	0.737699001068447	FALSE
d2ad8ffb9c564a7a919a6d2c2ea25bb4	8	0.737699001068447	FALSE
1b6acfec1f1d9d018036a1e275703706	8	0.737699001068447	FALSE
67fd3a3f3f8db748efbe268136dfbded	8	0.737699001068447	FALSE
4ce4c3a69273ce1cc8fd219da9b381d4	8	0.737699001068447	FALSE
11c23f6cbbdfd03bd8e6224a9071fb43	8	0.737699001068447	FALSE
f40c0d4d9e06093a5ada229490ba0568	8	0.737699001068447	FALSE
9ee2b7f0b9e9c362f67a773511bc0f8f	8	0.737699001068447	FALSE
6232643e6f00d30f14f28555f4d3aeb3	8	0.737699001068447	FALSE
bd51e684722285320e57f327eca15b6b	8	0.737699001068447	FALSE
0789caf27bf5974dcea4d8031951972a	8	0.737699001068447	FALSE
6d113989510223e4c6dbd981301f7212	8	0.737699001068447	FALSE
0522bfa94eba6ceddf15a808b6ae96b7	8	0.737699001068447	FALSE
ae12d30bf4100c4d96adc3ca43677698	8	0.737699001068447	FALSE
233979536096fdadbd5a84680c813451	8	0.737699001068447	FALSE
24df451d973741ec25daf4a0faabb381	8	0.737699001068447	FALSE
a31bd41aba32e4789d6da6371e8a0351	8	0.737699001068447	FALSE
4cf9490af23af672d19005d40819f32a	8	0.737699001068447	FALSE
d9681b85947e3afdfe94497149ef007f	8	0.737699001068447	FALSE
8817c9e2e994bf779d0d9871bfaab685	8	0.737699001068447	FALSE
fc2be12ee7778a775c5e96a7e119d6d4	8	0.737699001068447	FALSE
47601118850e3bfedf4909d626f8ac75	8	0.737699001068447	FALSE
5ee7c51bcfa27a4b318bfb90a7fb84c5	8	0.737699001068447	FALSE
5a300b9a353ac9001a36a58e3306a300	8	0.737699001068447	FALSE
e6552817f250085328087b23ab0dc1a9	8	0.737699001068447	FALSE
085bc0620ad4765efb14c77f0959dfef	8	0.737699001068447	FALSE
f30c82900b5992c0322e1142908bbe21	8	0.737699001068447	FALSE
1c12865114a7479a355a5314a0962ff4	8	0.737699001068447	FALSE
5608c3e6c9de9ceb79610e7786bd0ac4	455	0.73908199478054	FALSE
b39c338f5e964b6cb87e07f10badc6c4	26	0.746982706665468	FALSE
749e45bde9313270898b93b8c4becd90	26	0.746982706665468	FALSE
e33b20a92e8323c00165fd246aa9a5ce	26	0.746982706665468	FALSE
e9c0e61a0dcfb70f290be826a755c503	9	0.758208132672148	FALSE
b36996030f130f5113ee26664007dc10	9	0.758208132672148	FALSE
23a6b01030772ac2e83ffc6a9ec32215	9	0.758208132672148	FALSE
3e039130d1c5ce6ee3bf389b3e16a6c7	9	0.758208132672148	FALSE
3e78c2d4071900023caf060ad6e4aea3	9	0.758208132672148	FALSE
c3a4eb91dcbf4532380498b716445ae3	9	0.758208132672148	FALSE
680be3278c6b80b1b8d60bbfca191947	9	0.758208132672148	FALSE
dbd8c548553b3b93df3ac057d7d2cb56	9	0.758208132672148	FALSE
eb872b492619292ca3b24e42f436f99c	9	0.758208132672148	FALSE
fb8f2032bbe5e226156ce9778d2509af	9	0.758208132672148	FALSE
0c1c7ad5d0260ba6e3bdcc9590c3c2c8	9	0.758208132672148	FALSE
5322c1b4520bfe8d265dfc591df5a64d	9	0.758208132672148	FALSE
c3b153bdcbcbc594c78f50e103ff2b2d	9	0.758208132672148	FALSE
ad73f4ab37e8e5f7ab14dee5fdf5c5ae	9	0.758208132672148	FALSE
b2c3bb2ac840eba4aa8a9cd75f8bc4d6	9	0.758208132672148	FALSE
01080e9f6920e834dd4353e0006970be	9	0.758208132672148	FALSE
0b974216c890cb7d7f1b671e54adc25d	9	0.758208132672148	FALSE
b34dc93f446851e20b7664594eefea4a	9	0.758208132672148	FALSE
9422fdd782051bd355702c8bc5c811c7	9	0.758208132672148	FALSE
4ae9adef4e7cf64f39d31521ce36fe69	9	0.758208132672148	FALSE
6fc9ce36185f7f68f3a3e677c395f7fd	9	0.758208132672148	FALSE
dd1b7f0afcfca1df0f341790f1a2c1e1	9	0.758208132672148	FALSE
cc89995866a6e6077eea2beb57f8a512	9	0.758208132672148	FALSE
bf62225734b915ae3e034aaf5cab2fc2	9	0.758208132672148	FALSE
a51bfff8dc3d75cc03636e70bbfaf47b	9	0.758208132672148	FALSE
5281eb7b2a57c232332fddd223fe88f5	9	0.758208132672148	FALSE
0e7017c0ccd789c11023b2baaa6aab3b	9	0.758208132672148	FALSE
35bbc323229797ab3654d3eb5d1dcbb7	9	0.758208132672148	FALSE
50b02ed56e5bd7071bd7b387464838b7	9	0.758208132672148	FALSE
145042b121fb19b759c9a71cfd6c3b11	9	0.758208132672148	FALSE
070502fc07b0ab293ac6d73549d3b974	9	0.758208132672148	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

2d88b9122b9b887f026e2147867bc8df	9	0.758208132672148	FALSE
8a6ff32f51d7c2b5feb3a3c4507ddbcd	9	0.758208132672148	FALSE
f33382394a5a2e3263ec1b9462ea0b56	9	0.758208132672148	FALSE
393d85b4c2886b63cdb62e812d38b2af	9	0.758208132672148	FALSE
53d442e1e344ed5313bcc6bbd17496e2	9	0.758208132672148	FALSE
1e5faad8ce74f2172060dba245f52f96	9	0.758208132672148	FALSE
b944e446e87fa7a72fe2cccb95deee14	135	0.762399939224698	FALSE
31fafc31f41c40a6477d2bc148624981	27	0.762451888237852	FALSE
62ef06b414289af582c84fd276bdf945	27	0.762451888237852	FALSE
53be6d7e62277b36a2809a2301b96b9a	43	0.764801684642535	FALSE
23443f6140da99fd12fc5b0c4720a271	28	0.777097544572831	FALSE
da6a820d6236082bbee9da8395ecc162	28	0.777097544572831	FALSE
6f2414f3216c719e8f1a44de9f625c94	10	0.777152650112269	FALSE
5e4049b68851f2ee269dcac5f0f90322	10	0.777152650112269	FALSE
09b48cd4c1f3fc619ba1f437dae36dcb	10	0.777152650112269	FALSE
11e23ea5f65f90c78651bcbe86b74191	10	0.777152650112269	FALSE
4c457a70243c649760a1811351afeb0c	10	0.777152650112269	FALSE
257b9d2aab382808563d7f42d38d6964	10	0.777152650112269	FALSE
75fab318c01fbc0d769ab98f3d328f12	10	0.777152650112269	FALSE
f91246e65ef5f4a6eb9ffc444cba10f4	10	0.777152650112269	FALSE
0d004237311058005729686940671 9fa	10	0.777152650112269	FALSE
38347f000c351f6a8b58265d41aaa725	10	0.777152650112269	FALSE
91101303b2fa23ed605482c99f27db6b	10	0.777152650112269	FALSE
c8c110fa739e613f951fae4412d6f9f1	10	0.777152650112269	FALSE
126c4c4926f559db383ec0b8320bfac2	10	0.777152650112269	FALSE
e683f1888d6f1ab46f404be9a98d7a3e	10	0.777152650112269	FALSE
fcda020e2366b6eb9a4f922739f4e6ab	10	0.777152650112269	FALSE
42813a7ac2260688d1e5def012a3bf78	10	0.777152650112269	FALSE
bd51a40a24ebd30ef796a7d484fbcd55	10	0.777152650112269	FALSE
dccb9df56eb71cad4bff7ed8431a7552	10	0.777152650112269	FALSE
f0f33d61309d8999591d129526908ddd	10	0.777152650112269	FALSE
33cdd6119240b68f35d70c12fc0aafa6	10	0.777152650112269	FALSE
32480562e8e85d61074673393aa0ec7b	10	0.777152650112269	FALSE
47addf6684fda1d186def383b96fee3c	10	0.777152650112269	FALSE
74d46334ea0cf84cc845e9b24cacc55a	10	0.777152650112269	FALSE
c37fa601378354e8583bb4334eb4e041	10	0.777152650112269	FALSE
addf1e3d60f6910392223e40dcfa898b	10	0.777152650112269	FALSE
106c1bcf09642488d48e631dc007f515	10	0.777152650112269	FALSE
4c476bff9c90e545fcaf52abf2433741	10	0.777152650112269	FALSE
627857f1adf233e328b08b9f4ac9b055	10	0.777152650112269	FALSE
6f5da0517a69410b3b7702e321396052	10	0.777152650112269	FALSE
7fba832c4a81a2dc868d445e56d77ec9	10	0.777152650112269	FALSE
73e59f5f5a61052a6e7255dd9a08aeca	10	0.777152650112269	FALSE
7a21c9fea813dd4d7670e06c68be9aa7	10	0.777152650112269	FALSE
70b3e46a9f56d2957a5da8bbcc4108cc	10	0.777152650112269	FALSE
ad796ee030e064cce08108c72a91eee6	10	0.777152650112269	FALSE
1b156a3626c51ee68eafec90c33e0734	95	0.780667151943677	FALSE
b9d2c93750a12ff7981a77e82f6a1547	29	0.790952233328766	FALSE
4224ae75b61b4c9687df8ff56db506ee	29	0.790952233328766	FALSE
d63133c757e4c822d016db63aaa83f83	11	0.794648929648909	FALSE
3044fec0f8b0b1e86e093d156925713c	11	0.794648929648909	FALSE
d960f6eb68bf555c7af553d60d44c613	11	0.794648929648909	FALSE
8a3bcb66a63e718322795fb3cf5786ba	11	0.794648929648909	FALSE
f52e34a07ff2e814f50950ea181cb271	11	0.794648929648909	FALSE
0c97cf0c7f79074bda3693bd7391dfd4	11	0.794648929648909	FALSE
1e7e1e01adb7640a6a7d607cb61de9ea	11	0.794648929648909	FALSE
6728d5bea5db1509d2bc27d229d9065a	11	0.794648929648909	FALSE
c720ca5d5668c68cf397a7fc3dd7fc92	11	0.794648929648909	FALSE
7581960d6d42dd79720ae6185de9891a	11	0.794648929648909	FALSE
fd7410edb0e0122ad8c41680a0c8cf7f	11	0.794648929648909	FALSE
eb30fe79ff2790662309c17e86b78943	11	0.794648929648909	FALSE
d88a6c6c55c19b1066649b6d8e929f74	11	0.794648929648909	FALSE
7dd5364869a72d25f903f2a7fdd10feb	11	0.794648929648909	FALSE
1a8d886228576ed91c68865e528df502	11	0.794648929648909	FALSE
7e910a4671368a87833508c5b7e64b5c	11	0.794648929648909	FALSE
fe93a26042d60e3601d4b58c7ffc3a80	11	0.794648929648909	FALSE
36fa43c25ed6c2d07c29e03e2894c586	11	0.794648929648909	FALSE
2f6051215455021cba7fae075f514065	11	0.794648929648909	FALSE
88ad1024d096672da0ec59be76064479	11	0.794648929648909	FALSE
c2815e98bbc9a5c26391b06e88536412	11	0.794648929648909	FALSE
34572a2307da886cbce2b191be5f213a	11	0.794648929648909	FALSE
2f36633f10f651dce4741ab9073fbe57	11	0.794648929648909	FALSE
859cd1dcd7e3525daa1e111756201de7	11	0.794648929648909	FALSE
72f332a962517e2081df0ebac4ce0e1a	11	0.794648929648909	FALSE
7aa55291042b0dc85bf66204aeec5f3f	11	0.794648929648909	FALSE
3fb2f665e9e58a2e7ad1fdd1ee77a271	11	0.794648929648909	FALSE
9a99c0b7962e2d34c24e789e2581217f	11	0.794648929648909	FALSE
3ec4481028596ce277448c7cb5385a74	207	0.796503292277668	FALSE
3da914834c78256cc61d1c1c470e4552	30	0.804048399611448	FALSE
045bb03f6292d0b36cd77c9e6c5c4020	30	0.804048399611448	FALSE
196a053b371610ff9ee00b66df0b4d35	12	0.810804914479843	FALSE
cc351e8607640aefef66b8b1b35fb44c	12	0.810804914479843	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

60f0a102e1b3118286adc774264b5b7d	12	0.810804914479843	FALSE
88b3d853bf0ae026c108316bdbf7516d	12	0.810804914479843	FALSE
88f8c0389e4c0872fbee3232ebc0eb5d	12	0.810804914479843	FALSE
c2c99e10b5956b9c6b87131faf3a9467	12	0.810804914479843	FALSE
1590cb24d34f240f939466c1bca211e1	12	0.810804914479843	FALSE
3678d89e46e36d2dc9cdd1a84955bdb4	12	0.810804914479843	FALSE
a6ac3b4cd416b6f921f702db706730f4	12	0.810804914479843	FALSE
4d350a5bdf32548e11848a818277b8c2	12	0.810804914479843	FALSE
17473586f4b889c05139cc72c2e8062e	12	0.810804914479843	FALSE
db4c8dc0512e585410924ee31218577b	12	0.810804914479843	FALSE
e556ee5150ea74092458189357186622	12	0.810804914479843	FALSE
305ab46bc50af8c9b04a52005224274f	12	0.810804914479843	FALSE
6165cf874bd70cc64dd43f05eefe2264	12	0.810804914479843	FALSE
00a7b0037094adc4a8f638508824284b	12	0.810804914479843	FALSE
958362b80eb527e5d99cfc14727a45a8	12	0.810804914479843	FALSE
32891522d68b23b9fbbf71520146d04d	12	0.810804914479843	FALSE
b7dafc10c238527a63f20df8257221eb	12	0.810804914479843	FALSE
3be4093633164903309652a5ee96300f	12	0.810804914479843	FALSE
0ac81f5b1143565c25fdfb8deb4206e8	12	0.810804914479843	FALSE
2d30c2c53fd94f63d88b136d45aca3b0	12	0.810804914479843	FALSE
0251c873b9e6e76973408b336adeb8b2	12	0.810804914479843	FALSE
43bafcbff59dcf1ac8758da64ef9b5bc	31	0.816418212900771	FALSE
8d5e9411bb401bbf42175d0fd956d125	31	0.816418212900771	FALSE
3861ee62617a071584eabec1162b5126	31	0.816418212900771	FALSE
1b29cef84de6cbd9cd52eb9dbaa7fe7d	31	0.816418212900771	FALSE
9b9a1ae38e23758a979eed98d5f06dc9	13	0.825720709603839	FALSE
e7de984e21341b13ac754739096343b2	13	0.825720709603839	FALSE
e95ab89ff87282bb7ebfb109bb235db9	13	0.825720709603839	FALSE
7a17d452d41caf71ea4de6188fdff66e	13	0.825720709603839	FALSE
a0f355228c2bc92f7151668f1fc71558	13	0.825720709603839	FALSE
581add769634c1bb41fe0eac84da9423	13	0.825720709603839	FALSE
a6e12567a5a8ff604652fed640545653	13	0.825720709603839	FALSE
047c596149ab5e1e62eb9341525b0276	13	0.825720709603839	FALSE
a8146cf9c7c06aec25a563142d473b10	13	0.825720709603839	FALSE
4b36eec432ef4ffb3464a8a0a9eafee3	13	0.825720709603839	FALSE
5d55c5aaab36a98540bcd71cf23696b1	13	0.825720709603839	FALSE
c606ef358bc6506aca4c4b89cc56e45e	13	0.825720709603839	FALSE
6fdb0a43d61c08d2d4e244d97a48e9f2	13	0.825720709603839	FALSE
91702465b934dd02c76f76cfdabdf946	13	0.825720709603839	FALSE
997eb2712b5cb0cec6f883c8f2fb31b0	13	0.825720709603839	FALSE
92c64b7632c36c15c97decd4597f4f72	13	0.825720709603839	FALSE
475abf5e533ed261100a749ba2b5f8ec	13	0.825720709603839	FALSE
34c94d66644c9e3e641a5040c3113f72	13	0.825720709603839	FALSE
a9266e9cb66778a8c6c548877d9e204b	13	0.825720709603839	FALSE
c3018532c596c385e1f4cd12b2d3fe8d	13	0.825720709603839	FALSE
52dcfdea4e1a3fe11524ecda68df475d	13	0.825720709603839	FALSE
55166f0153519210b5f9ee37f47f518d	13	0.825720709603839	FALSE
4f10e321d38c10ef2b7a144f41da7d91	13	0.825720709603839	FALSE
edb6db65965a970c9eab814177f21796	13	0.825720709603839	FALSE
be6838d85b6c9b7429f1f7f4879c1d28	13	0.825720709603839	FALSE
f3325b6c39cd939584b540aa9894f1d5	13	0.825720709603839	FALSE
e19ce815c8643b851715aba05acf960b	72	0.835326575452095	FALSE
243efcb3cb4808110c65c8176347c6ea	33	0.839105266188764	FALSE
b98f8c87cd3159b66ec0a418919be624	14	0.839489135872143	FALSE
4d02c00f27d2b0369cae8697f8aa74e0	14	0.839489135872143	FALSE
804da4953d5a3c2dec62eceaafd77dbb	14	0.839489135872143	FALSE
e54de8f948e4f1d2a3c9e5a394c5bc37	14	0.839489135872143	FALSE
7c664e537af176112682a9792e06e757	14	0.839489135872143	FALSE
dbb77ef1903ce7b239d6c65a843ad3f3	14	0.839489135872143	FALSE
7d8a38181893e5747a2b6c5a1be1b9d3	14	0.839489135872143	FALSE
d9d7d7b8bf24da8102e83469458682bb	14	0.839489135872143	FALSE
6f3994ea8e29e64aa02df14a2846a8f6	14	0.839489135872143	FALSE
5655eff59411150b46cf6f2d75b35308	14	0.839489135872143	FALSE
f30cd40fb114ab501e51f2f8aa202e14	14	0.839489135872143	FALSE
b5df6b3e09aba5b4c07aaa7690198885	14	0.839489135872143	FALSE
670ab9f3126df45d2280c6d68eb6724d	14	0.839489135872143	FALSE
61e1d16b27c60bdff7b9d4a61877b777	14	0.839489135872143	FALSE
46bd431314d551a4394037b4b0e5f9da	14	0.839489135872143	FALSE
18430608bf7e39c8866f4592dae12694	14	0.839489135872143	FALSE
8f2f800e4a8e5c41bed76dbcb81203b8	14	0.839489135872143	FALSE
d5a541cd845afa61ac26e362a9368819	14	0.839489135872143	FALSE
6b32fa82cd34029494b5351fe148c7f7	146	0.843871452255949	FALSE
ff4ae0f811e8dcb2d76d7d8acc73deb8	132	0.844381485294231	FALSE
23d922f2b3e9315c90bb4189ba7f7f52	34	0.849484318274827	FALSE
621a423823ae9010632f322b6172946f	34	0.849484318274827	FALSE
6b11f2c5e75502f47952b01fc902ce16	15	0.852196245948932	FALSE
aa806b10a642f7016ecf85fa68d73b48	15	0.852196245948932	FALSE
255e84a7816e3f4f5e5c82c5b086195d	15	0.852196245948932	FALSE
ba9601da497863819d4758229b591e4c	15	0.852196245948932	FALSE
2c894c9e3a032a9824774aed6a96603d	15	0.852196245948932	FALSE
8d9f970561fe1a7c174364c656380fe2	15	0.852196245948932	FALSE
2bab1b7d627ef1fb38400f422073d46e	15	0.852196245948932	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

ccd93a6f7c86fa38459587813126e9fe	15	0.852196245948932	FALSE
15b8ef9c8fd1c209fe10fe79b57a36f4	15	0.852196245948932	FALSE
7cf993146775b44a6b8f89833e71508b	15	0.852196245948932	FALSE
8d4b1c33aee9458da55ed0181ca9b8e3	15	0.852196245948932	FALSE
882f7684e960b257722e753254c34447	15	0.852196245948932	FALSE
465a1ce875ea713594915ac415e88ab9	15	0.852196245948932	FALSE
ece50a62168f85fc61385d8adb4c6494	15	0.852196245948932	FALSE
8a185c2d22cabacdf694587e629c5e53	15	0.852196245948932	FALSE
6b952a73de08f9fe604db1f30be44133	15	0.852196245948932	FALSE
7227f56c77a1a528f4eaa76c94e2004d	15	0.852196245948932	FALSE
4eb188c973c9d35a8ee1c68a9adb9a42	15	0.852196245948932	FALSE
a2d516468a24a0b51943203260a1e7fe	52	0.858720990665559	FALSE
ca79aa0ff5cff05327203d05dc798f8f	35	0.859260460867701	FALSE
e85924823aec7bd0bf33576e477e9902	16	0.863921804725426	FALSE
0e5ef2615e277c7143361c56e09c0b6f	16	0.863921804725426	FALSE
f33a119e73a383eef2ba3f02c6a134de	16	0.863921804725426	FALSE
65d47248136b94e72d9aa1fb5392b608	16	0.863921804725426	FALSE
0bf4f18b63b6282536d07bb07e9a993f	16	0.863921804725426	FALSE
82ac56ca828692be621d100e94a3d6ff	16	0.863921804725426	FALSE
d0ee179b720c3f1df4b1b44f35471a75	16	0.863921804725426	FALSE
a29e586ce34cb2edac8c292267d77ad2	16	0.863921804725426	FALSE
b8a62348b7c542b4e7f94bf0ccb486d6	16	0.863921804725426	FALSE
5160cc3e958946a0a4337b33b6533d3d	16	0.863921804725426	FALSE
88f0846ea1250a0029afddf977ba0e5d	16	0.863921804725426	FALSE
97d4e9ea1f067927134c5a34f0d5ad06	16	0.863921804725426	FALSE
a5b07cf8520236c817752cda1c893a9c	16	0.863921804725426	FALSE
368c8aaf10922a92b5d77aa9eb36a7c0	36	0.868462790609597	FALSE
a9a8ba05b0fcdbad9ee1282d3076b726	36	0.868462790609597	FALSE
28b378e38018bbf0b0484af036ef1e5d	54	0.874611180745984	FALSE
c8fb6ebbb3507d5257309f12e5e4f705	17	0.874739736567337	FALSE
fcff23401e2280bde7682f18fa1d1cf7	17	0.874739736567337	FALSE
60cdff71a273a431942ece7f772dcb10	17	0.874739736567337	FALSE
de4efdb3a52a4d915422b3bd60610041	17	0.874739736567337	FALSE
18c8488af7fada9ef923d6452aeb3a9d	17	0.874739736567337	FALSE
df2dacf39bb2269336cd04a5fefdb119	17	0.874739736567337	FALSE
c184aa75193f4e1f2a583a1fa827b3dc	17	0.874739736567337	FALSE
18b30b18da1d03ffe463fc69ef6441c2	17	0.874739736567337	FALSE
02e3c541b67c1f742460bbf565f63071	17	0.874739736567337	FALSE
1469f96bc2b17e99a7988e3fc891607d	17	0.874739736567337	FALSE
56b22d768a1396ed9e442d54a5ea4dbd	37	0.877119570694495	FALSE
a4d01993784ebbe3fec3839448d1f52d	94	0.87993196455902	FALSE
af41b8a7c03317c74926ead0fd355e8f	18	0.884718541620673	FALSE
73858fe4e97c180d5a7545ff94a6a788	18	0.884718541620673	FALSE
7fb82c9bf8a0918e18e5fd25a3064d11	18	0.884718541620673	FALSE
2d95ca79269a5201b9100b5bb7f969c5	18	0.884718541620673	FALSE
c3d3ad1283cfc17a3d466847de0f585c	18	0.884718541620673	FALSE
642e62255f6c842272cd37f56c189bb6	18	0.884718541620673	FALSE
e90425a6fe76084e079af542f4bd2632	18	0.884718541620673	FALSE
6f34eaefa706c2a148d34d4b640db857	18	0.884718541620673	FALSE
2895e2dc704bf22b9ac315e117faef1e	18	0.884718541620673	FALSE
06957346beaf0bfdf292900cf66ebdb6	18	0.884718541620673	FALSE
073718ac989b6696918a078173a59117	18	0.884718541620673	FALSE
43d98b132b922ada4a8b34d08dff7d38	18	0.884718541620673	FALSE
2d6f7915a01751f1ab279a4e17c15b8e	18	0.884718541620673	FALSE
8ac40970733480af91b7d4204a5ee421	18	0.884718541620673	FALSE
43f2b9bb3ec86a9c4b2322f3b1ca9660	18	0.884718541620673	FALSE
24513f3439ddfae0170dad90ebce403b	38	0.885258189454661	FALSE
df45ff114fca650e2150e761e33d947f	56	0.888943713616188	FALSE
e7929c7247681ff65b78376206017180	39	0.892905129428559	FALSE
a48c86bb43aa39dce41597a3ff0a026f	19	0.893921683255998	FALSE
794184e730540cc7ff89c9e06d29a1d3	19	0.893921683255998	FALSE
fe44c3d1e39fa07e328e63b858248557	19	0.893921683255998	FALSE
6c051d317e4e22c0ed7e6a0f1d665337	19	0.893921683255998	FALSE
8603c3db2bb660bfd3b8d633a4bd4c7a	19	0.893921683255998	FALSE
fdf86a1fa9f0935a6512bd850d5997ea	19	0.893921683255998	FALSE
9370d8d53940a56f66653b13014c2652	19	0.893921683255998	FALSE
413bcb036aa08e6c887c47488e63cfd7	19	0.893921683255998	FALSE
b448d68001582a65eae812c8015a15d2	19	0.893921683255998	FALSE
cff0ae11f6b5cd20ab0d14c2cb6f773c	19	0.893921683255998	FALSE
a9d9e22d46fd432964b7e70ceac35fd0	65	0.894119310001397	FALSE
25089d3f7bed37ea7e764b89a1d2f4e0	40	0.900085945646662	FALSE
cfdd70bdbf090e219bf771ede00b81bb	198	0.900672614915809	FALSE
56de38318fb195cc98890d09d1fb9fd7	20	0.902407948595518	FALSE
c3ecf8e6e7e7a856fbd2df2d4ee56b56	20	0.902407948595518	FALSE
5797b86035ec7c8fa28465a07b9805f0	20	0.902407948595518	FALSE
0ddcce08d539797d98bc25c7c02ae7f8	20	0.902407948595518	FALSE
f64861919bd8bf28858e81c625781fc2	20	0.902407948595518	FALSE
46f6ee1a78f2fa98ad1e29c553d32636	20	0.902407948595518	FALSE
1276c14169f9d464cdb5471ea52e902d	20	0.902407948595518	FALSE
ff2bd29ff42e4dc25a31714e0b6c2dca	41	0.906825251993826	FALSE
5f3673305f1d86e0ddbfa1ed980368b7	41	0.906825251993826	FALSE
4119e9aeb272d4ff8f0af3a026865785	59	0.907777302682907	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

8db9ec3c6c2da03db22d64f72b6f0c7d	21	0.910231783940181	FALSE
6ac028dc5e7f0066b35194e99b072aa8	21	0.910231783940181	FALSE
a1bed12a320d811c7025e11b4aa933ec	21	0.910231783940181	FALSE
d0e7ee826a8a40ab80d2bcc61ab571f5	21	0.910231783940181	FALSE
0a1561a33d1b52b9b8846598ed4ca557	21	0.910231783940181	FALSE
8fd632a60760954f8497f1fc8a51f4b3	21	0.910231783940181	FALSE
219122904013b99064a9d2cf96b0106d	21	0.910231783940181	FALSE
f48e415e85476e0bc78c073767ad7196	21	0.910231783940181	FALSE
ad460170fa30ba7e8f74c918c6fbb3b9	21	0.910231783940181	FALSE
eb19b3f753e323f3d3d73a77b517e46e	21	0.910231783940181	FALSE
ee4b74373fc0cece03796e5db2aa3d3a	21	0.910231783940181	FALSE
2cab47f2bc2f1bdaefbd331f3a490414	21	0.910231783940181	FALSE
63355ac2e4dd546d908aa62e078e814b	22	0.917443606794881	FALSE
c324721a2dc34076ead3590c735a6779	22	0.917443606794881	FALSE
af8c2f69e1dece5ad3a2f23731ae5420	22	0.917443606794881	FALSE
aea976fc75991164def44c3dc183bb16	22	0.917443606794881	FALSE
cd9caf8d961e5c6a16348abd971aea04	22	0.917443606794881	FALSE
a5189f77a2cfeab3bc1602ff5c8ac3e9	492	0.923076923076923	FALSE
31f23810ce7087481ef862d15d9fd041	23	0.924090096078344	FALSE
3d535665aaaee9491f84783ece3dd889	23	0.924090096078344	FALSE
03b3b240d243a428811ea34a257d4bd6	23	0.924090096078344	FALSE
5cda16c50b985a14e0ec9910e3726133	23	0.924090096078344	FALSE
ec23b923e536e0a2279f4c973531e122	23	0.924090096078344	FALSE
e952ac57af197a320c22ecec62384f1c	23	0.924090096078344	FALSE
88da24edb3fb401aa506d44503445f2d	23	0.924090096078344	FALSE
2d834dae498cec30f6b9043cf2910a9d	23	0.924090096078344	FALSE
752353b9793ab0e9cd4dab60c7b72860	23	0.924090096078344	FALSE
bb00e9f32a7ad532a6c3ed50dcb3d251	23	0.924090096078344	FALSE
1afda3ce15867b1740804995147a56d1	23	0.924090096078344	FALSE
94202d625126dcb1729abbd7dd7018a9	23	0.924090096078344	FALSE
cfa03a7d8244ddf2a84841f464810ed5	23	0.924090096078344	FALSE
a67f11c0453dc4e295b206418d36b1f6	44	0.92462603707359	FALSE
19b699fcf40dfdeb5ed8ace52f4f2dc8	44	0.92462603707359	FALSE
d1566cbf914a151f452e881fc9cae2e6	88	0.92543879210039	FALSE
df72998479c8b1228c5c3943f047680b	177	0.926350766924639	FALSE
bba4e5e8793fe193ffacb3e7a66c9f29	63	0.928502518965154	FALSE
025f3ec3dafc2f35bc840c5669bccaff	63	0.928502518965154	FALSE
15d1ae7f19a3b5bcf4667544bde11d38	45	0.929826511957183	FALSE
4cf3fce832b0e75aa582596e5159e808	24	0.930214461999836	FALSE
5df57ffc227fabd2e10b3d6652ccc022	24	0.930214461999836	FALSE
1c6fdc27b561532077e6d9e80823a7fa	24	0.930214461999836	FALSE
e1b3da5566b52686c72b1c8b37d79ab9	24	0.930214461999836	FALSE
3db627cc5e513aef2a1a4690e49e095f	24	0.930214461999836	FALSE
54f785f2d739fa4cdf5fac6da3ebfe4b	24	0.930214461999836	FALSE
22be50f25335287f88a3b9c75961865b	24	0.930214461999836	FALSE
1c0b93dba78bead886310532c780520c	24	0.930214461999836	FALSE
4bf24b715c00092ece6c83803f8f80fc	24	0.930214461999836	FALSE
bc880c65f6166950172bc6317df770e3	24	0.930214461999836	FALSE
c652553c5dcf87deed14605ad770aef4	24	0.930214461999836	FALSE
deba196a20f64e12587b394043745e35	24	0.930214461999836	FALSE
70f27b663b51acdefb4c6e059612bcc5	24	0.930214461999836	FALSE
f248f33104991f8f21e7cd8053710b68	73	0.930864893815069	FALSE
28bd824140d5c67d97ab2272d1014303	74	0.934524190979015	FALSE
0e22b2e49a15e74e09bf1c151ed0f16a	46	0.934694395790298	FALSE
f8becf0c50c828245f68b0f69562ad0b	25	0.935856696987084	FALSE
e3e4029653319068311462 81f74a361c	25	0.935856696987084	FALSE
96716aeeb18233697ce5131e6a0e778a	25	0.935856696987084	FALSE
077e11092881c0e616329dbdb02051ea	25	0.935856696987084	FALSE
f6ecf45f639b721f6645c142b25b91b9	47	0.939248703899328	FALSE
5ee5bf5ebc1547aea9d338e88b0b7c3b	47	0.939248703899328	FALSE
642567877ebcb1fc9c9b26a7ea85fbc3	26	0.941053808958279	FALSE
7acfbffd36ba9fa46063bf2688c97afa	26	0.941053808958279	FALSE
bae9e2ba2c9f33c59e2c4d6127de0a3a	26	0.941053808958279	FALSE
e48e37a0467ce3d2ff2c3a0f1167497b	26	0.941053808958279	FALSE
2c22a0fda496e652c971a6cfec88ec26	26	0.941053808958279	FALSE
8b1f185c948b07011f747c2da9878f91	125	0.945011945405258	FALSE
75d434e8cbac96a34057771b55101151	27	0.945840038145428	FALSE
16458c9597c205072815592a80ecd7c5	27	0.945840038145428	FALSE
b3bd84d7babbb68dbd7a1528593265c4	27	0.945840038145428	FALSE
70f04406961b7b38ab74ba06cac3c090	27	0.945840038145428	FALSE
933bef6fc914f6f30ce2edd6cb048102	27	0.945840038145428	FALSE
d4dc3f83e8ada91f63124bd347103ae9	28	0.950247058596121	FALSE
5a4eda6fcdd6ab834dba0111fbe55d95	28	0.950247058596121	FALSE
1a5c35a922fd259929f1379e9fc89fe9	28	0.950247058596121	FALSE
61324ada608d3a8fc4f97057a7acd186	28	0.950247058596121	FALSE
4bb0efcbe14db9a98578c065540b3ce0	28	0.950247058596121	FALSE
d453b2abebc22f909d8e87353f4d03ad	28	0.950247058596121	FALSE
c33766d386623db8e71db4e9d44a2f8b	28	0.950247058596121	FALSE
a1b5c85117491a6fee703b82718600c6	28	0.950247058596121	FALSE
6cd8021d0df25c00c8f825e988b87b40	79	0.950296797263479	FALSE
6091b6f3a98a65dea5b6add6b3b7d887	50	0.951207180480497	FALSE
a73619885909abe32871924ea5d5573d	50	0.951207180480497	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

8e7e470f27ee9dee24fe1063312db221	50	0.951207180480497	FALSE
70aa5a01ed7683fd122c10659d43706b	80	0.952997495063912	FALSE
1a95e5b78913fedee75e1f052399e016	80	0.952997495063912	FALSE
9ba80d636e5561c9a76f1aa0f67b5b4f	128	0.953018111878516	FALSE
dab9247c26155029ce7821e765ebc155	29	0.954304165405879	FALSE
89c4dd7f04661867c45fbdc395b3fcee	29	0.954304165405879	FALSE
480a9d52f57963fd090a9dd94ee44681	51	0.954679968359587	FALSE
07bf7669e42ebc17a676ae314133677b	114	0.956449115914962	FALSE
c7985a28aa4cf9627133974bee02eb07	30	0.958038448663033	FALSE
c4bffd0ab73cc7a0273cee68fcf49e24	30	0.958038448663033	FALSE
e27680d4009f98f30248d823bc17fb8e	481	0.958468877945874	FALSE
e71f2cd748d4be5bf9412164728834c2	65	0.95978801932764	FALSE
8d278be881c77f270f9250ee0f55ff8a	437	0.960667235356931	FALSE
8ba130bd59fb89ff9f115561c5945f96	53	0.960945634967021	FALSE
c88aa34a8ddb9b4f44cd46b7ca1ebdd4	31	0.961474955022526	FALSE
a2d411fff3a418ad531287e0f24eb090	31	0.961474955022526	FALSE
fbdd0ce9feb0e5bc7494354ad9296b92	31	0.961474955022526	FALSE
d684d6466b2c6e9ac91ecc13f105b1c4	31	0.961474955022526	FALSE
c0a5a04e7e5773e91892d4c41ab7e0cb	84	0.962505104355992	FALSE
2d983b4684cc85598f7454c72b58eb26	54	0.963765892214911	FALSE
9a2dc4f076b397af87aa6a5cf4404d41	133	0.96413379708232	FALSE
c5803199ff4c3da81551b3b5223ec5e7	67	0.964214349147749	FALSE
b519b2fa0b4b23dbc16576634c819d16	67	0.964214349147749	FALSE
9bee9198d6fd38b817cf2f3a7b118763	67	0.964214349147749	FALSE
23951d86554a06257abfffe1b4bcdf8f	85	0.964587351118251	FALSE
ac181dd33b4fc009cce071af5f0d21c0	32	0.964636837763658	FALSE
e9fac19935b286ddf2a78075e7017977	32	0.964636837763658	FALSE
3d8c220d0ff87025f87b4955bce98e2b	55	0.966394818614568	FALSE
664ba36a3ef364511b4b5865f662ea58	33	0.967545496129418	FALSE
40569c7ad9644afd194ec2f0c317d157	33	0.967545496129418	FALSE
38c2834f2d440309fa6dd0bdce32ed78	33	0.967545496129418	FALSE
b05bfcc2ed94261623c92866941caae4	33	0.967545496129418	FALSE
8c29c0ac4bed029ebbeb22b3935cc8bb	69	0.968175996315455	FALSE
b28ed22c1484414affbb2110e2cb8dd0	70	0.969997389487643	FALSE
73df01709999f4490d3dcd729cb9665f	88	0.970211815168778	FALSE
502d89eeb4d74d570458302627aa1beb	88	0.970211815168778	FALSE
19b611714457a10b79396b8ee7dfbba5	34	0.970220704691418	FALSE
34e6dd3019805dcef9989972c300b776	34	0.970220704691418	FALSE
6a316c83cc725ab296f381e1977fcc84	34	0.970220704691418	FALSE
e8760ab74d8bf916e7b9e6742d3fb1fd	34	0.970220704691418	FALSE
840600f831582858ed2ae3964571e4e7	34	0.970220704691418	FALSE
05a9bdb05e34d5848bf5144f0e58419b	380	0.971009992641975	FALSE
57d52f46acd5562d4edee17e354f9cc6	57	0.971125753566493	FALSE
d3247c936f8f4735909a8526ebf2f49e	391	0.971435119747249	FALSE
ae4ab154cb983f0fccdc437f469417cd	35	0.972680733434301	FALSE
b2bbfa9e2ea7dfdeb3807096c69b1e9a	35	0.972680733434301	FALSE
fd1c43ccf4c93ee1cd46c142e1485689	35	0.972680733434301	FALSE
8c623c0f9544735a038e1d30d2b51fa2	35	0.972680733434301	FALSE
d4f532a400b5a2e27e030558ec7e3a10	35	0.972680733434301	FALSE
ace63ae9de78bed2e5510a38c85812f1	35	0.972680733434301	FALSE
4c06188bc9c7b91eec9e1fe25da00706	312	0.972895070666662	FALSE
fc94161516cf5edaeb8c984c59c9ab21	73	0.974888026979453	FALSE
7ce80a763e15761f1fe86ef8581f826a	36	0.974942459206624	FALSE
a347935cc17b25761d608e0a04fa70ef	91	0.975001953706185	FALSE
ca461f4b5e8bd616a405daf9f2ae0f20	59	0.975226221953246	FALSE
d6be45afca6371cdd7d012d8ce94ebd8	108	0.97523648938085	FALSE
ba673c958da8e9385196463c6ef87a4a	140	0.97582220885328	FALSE
fa5757ea2e0bd2ada15a5592ae37c723	37	0.977021469141446	FALSE
51700328b249bef1561f6d7115121c77	37	0.977021469141446	FALSE
52f7717d7d4b29e0a22c4e08ed966ea3	37	0.977021469141446	FALSE
48cd0867fbe39c565a7343262fa9585e	37	0.977021469141446	FALSE
e3aa1e69ab25c04321fca9f0996fb994	37	0.977021469141446	FALSE
446d2c96ce6aaa50144d2da771ef8863	37	0.977021469141446	FALSE
e8e69f28dc2939bc04a1690ab27bd4bc	75	0.977718738157169	FALSE
9dfd78e79bba89ffe82035671fd61f9d	61	0.97877442548296	FALSE
60de5a18f62043a523b08ecc4e10a441	61	0.97877442548296	FALSE
0fb527bad6c2916e89a972cf6eaf8a42	279	0.978835312501384	FALSE
48957f322815bd853ff040b697d46043	38	0.978932156608897	FALSE
1320f4cb4f2f04a8f6c7709b994b937b	38	0.978932156608897	FALSE
ad65b63c122ba12caca48384f0f34ddf	38	0.978932156608897	FALSE
28323331d787c7344fc7407d7e26effc	38	0.978932156608897	FALSE
8134209f8d6001f476f316447084e099	77	0.980246285334548	FALSE
f9ff9d5eb1c59d26a9689d1b617eba2b	62	0.980363455186265	FALSE
66e6cf9c166194e84fdcf94637ce3bfe	39	0.980687810224822	FALSE
897b61e5d2dce1bf89fedb73ed0eca90	39	0.980687810224822	FALSE
45a68a9eee3cf83e27f4ea309d57ffc3	39	0.980687810224822	FALSE
3b150d94f8782944249fd0a73999227e	39	0.980687810224822	FALSE
5a8154c58ba9ea7bdadbcff21b3a3e6a	144	0.980868578285112	FALSE
8d16916dcdaa69cee304e39539b62bec	40	0.982300696403848	FALSE
ae6ea074d9cd8ec6896004a16d81702b	40	0.982300696403848	FALSE
0345b451b2a01dbc32e694d5310d02aa	97	0.982522440946451	FALSE
501dd344beeb7c69607fa1ffdf9cdfdf	41	0.983782135911896	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

d70bd64edc0672b74dee70fad8437752	81	0.984512697502304	FALSE
c224c0288f2ed4a3bc2220986cad1932	42	0.985142574842024	FALSE
6d43cac09b0febf5f95abf46b0080739	149	0.985852356254384	FALSE
5f7b1039e8789ade0b68ff96715ebdc1	43	0.986391650408402	FALSE
8cbc77e3bd6a3cacd92289ed7497eb49	43	0.986391650408402	FALSE
7020e0243ea7e1d111d3eb43591f550f	43	0.986391650408402	FALSE
15b888fac2a8be4df2917b6670dad73f	235	0.986912054992386	FALSE
53dc126c6dcdbed4cdd3eecb29bbc2f5	84	0.98712561921398	FALSE
4cd36d8f6046c25842ef887d47e7b023	44	0.987538251926098	FALSE
67acfdff1a08a501e960ccc83a5c62c9	44	0.987538251926098	FALSE
b08ef9b722e924b7d7758709ea999bac	85	0.987900140202004	FALSE
2e94d4dfa19d87774d333b853fcaff60	66	0.988438983428952	FALSE
0339d50219c71731139f9ce51ded15c6	66	0.988438983428952	FALSE
3ac305e8c1b4c9ea90f35d8600f8b8d6	45	0.988590577319005	FALSE
830c3ea3526c42f09986097a09800338	104	0.988632991898301	FALSE
1e8a8c7842c164cb6f6c6029625e902c	67	0.989132559179371	FALSE
f924b59e41bdb872246acf7db5361fc2	67	0.989132559179371	FALSE
d3640519d63b009a839d93bb86a6bb29	87	0.98931962583368	FALSE
6b72125ffd5c3d74a2394b56bb6ea404	105	0.989322858710674	FALSE
8643001f205ade41838a964499dd97b8	169	0.989648763831626	FALSE
e2f82f601e2f4d6f8cc9001faa84d7e2	88	0.989969187801586	FALSE
bd23f3464dbee38f2352549d726766d1	88	0.989969187801586	FALSE
fb9868c77db10fc39a0129cde8a6d9b7	88	0.989969187801586	FALSE
52433c56fb7fbc1f94822c886b18b92f	69	0.990402291856229	FALSE
8dddef875ae1fd71b2f55cee552f56e9	47	0.990442044742412	FALSE
d2b22a6509e29fd8d160f8af18373fb7	47	0.990442044742412	FALSE
b86810a3acc42fe737e424e186df83f3	47	0.990442044742412	FALSE
f3b9cecc06cf0c99e801d491b67de2cc	89	0.990581499379893	FALSE
60efc1e5aaedaa4974a098acf1a85a2b	124	0.990649125722333	FALSE
8f98fb8693ed59c21399d83ce2d10724	156	0.990888861088775	FALSE
62bcf76a05df8d7152251dc1abacc559	70	0.990982774941752	FALSE
a54623313df876441f6bd6ea22292102	48	0.991254577851558	FALSE
4fb4479ec3e40df5abc4b828e8a43e85	48	0.991254577851558	FALSE
dd8f88921af5ba9b0e334e9b08aafe18	48	0.991254577851558	FALSE
8af0ba30611d668e5a3927f85542f627	109	0.991712936225779	FALSE
a86b3fc7197f2a1363443bf2934eff4c	109	0.991712936225779	FALSE
55af53a70bb5d746e81b9ce3713a110f	49	0.991999703505461	FALSE
2448d5a801ab0985c09a7c265d15308d	49	0.991999703505461	FALSE
e27762562e5f097f5dd5d3ddaade6200	110	0.99222751881737	FALSE
5d564d5db4ddb67c61b7fdfbffbe15d3	127	0.992284862432369	FALSE
1ff5d9c73c320602c29d8f6f67907114	73	0.992530200713038	FALSE
4ade57e5bf1dea163c998a806258ee84	50	0.992682874891512	FALSE
b4204990d266b67c13e644d5e2e3be91	50	0.992682874891512	FALSE
dd4411cbe45e5040a3a75c1c10cfa747	74	0.992987246573161	FALSE
c991e7b4a3dee5ea8515b854f6631824	74	0.992987246573161	FALSE
9e821a0ea32dc0a79c624893fa1742aa	145	0.993189977710979	FALSE
b121a17a2f7b359de8888212613f319c	129	0.993225110763305	FALSE
a75fdbd0512a2e0b09b67946674ea28e	51	0.993309115328724	FALSE
3ff22a19150bb5bb3fac504b21dd1e34	51	0.993309115328724	FALSE
1860bcc07a862435fe3fc78797201dad	51	0.993309115328724	FALSE
573610ba002103484c779191832a6c9f	76	0.99382283300343	FALSE
51e441cbdcc80da0656e82293ae160b5	52	0.993883051259895	FALSE
96a4b5621edd6a80e9bd93ec87b70e47	96	0.993981542524009	FALSE
e7610857aa92ccaca28d9abdee16c246	53	0.994408942780537	FALSE
2987f11421aba677b3ed1ba21faf8907	53	0.994408942780537	FALSE
b63b4aff4d8c479995ebbb5c206efcc5	53	0.994408942780537	FALSE
c45aec1dff493b3b3a92b769c3527575	53	0.994408942780537	FALSE
e5698bbc0af657f3e43e3cfbee8f55c7	53	0.994408942780537	FALSE
731043e4dbf3a367cdaae8b91bdddbfe	55	0.995331968584332	FALSE
47eacf9c2676422e360170281752eba7	55	0.995331968584332	FALSE
f033e2deafb6fd1525fc91a4d76a8b1f	100	0.995366986351095	FALSE
5a722836ab0e0124d64413717949ade4	81	0.995518134373802	FALSE
686b1c1752a35be01421dfdacfa244e1	101	0.995663257771153	FALSE
b1164c7556428ca28a0fa47436b70504	56	0.995736035085005	FALSE
d5a1283783f4f7170bdef0cd036ee48f	56	0.995736035085005	FALSE
ff7520f4d531ff12bb529c5403dbce5f	82	0.995799422473728	FALSE
9ad50086ebb6725181556774a235e206	102	0.995941708331355	FALSE
74f7929cfd22b1eae037d866281cfc63	83	0.996063945953789	FALSE
118ec7a179026c54887b0f117b8270b8	83	0.996063945953789	FALSE
a9b2c08281a57e2decf34d0892ba7947	83	0.996063945953789	FALSE
a3d81bc3894304b7c0a6df55297a8d08	57	0.99610596811416	FALSE
37ffc50b19370c0b659f4c3d919f99e8	57	0.99610596811416	FALSE
9dfda2206e7dcc1438c2594ff36f4b94	57	0.99610596811416	FALSE
96bf4520c8e056750f57133a13ed76dd	57	0.99610596811416	FALSE
cacfa81f2bc89116d93fe225248244da	85	0.996546457471862	FALSE
5aa73bdfc0eca9a7c83f64620d8828f4	59	0.996754455673928	FALSE
4f27892b2ff94c7f0d0cf982f27d8541	86	0.996766197016136	FALSE
56fd8c08e9e106325cabe0a193e11914	106	0.996896639309482	FALSE
58238c3d1af33f7a1e7c6769b3870a66	60	0.997037974488619	FALSE
a1c5c150d75a73a1e90ce9724214d0e3	61	0.997297322344454	FALSE
35e06063493e115e43f15cc0d438eab2	61	0.997297322344454	FALSE
53805a3bc578a4a6d9b8dcb1946e6d76	89	0.997348877863528	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

1643fdafc7ede484acb8d727168ab741	189	0.997385663338198	FALSE
e6fc000a27af86cd26973287af4e061c	109	0.997469994209178	FALSE
5276610a51db67a6a227a6a3b486efd0	90	0.997519988992907	FALSE
374a0952cc1222bfe7e3f87423eb96a9	62	0.997534508836165	FALSE
8b4400776aa5b6f0268025915d1bee5a	62	0.997534508836165	FALSE
f4ff94f14b1feb5a48cc311c5c895ea0	62	0.997534508836165	FALSE
a8350956db53b9cb8e376f9406f154e3	218	0.997537442777406	FALSE
9ee696aaabfe93f028faf08377028b06	110	0.997637937380995	FALSE
54e7b2139640ec28fbaf7f0491cad845	91	0.99768064271719	FALSE
a26ea2de067690b7f7ed5e1d1efa4770	63	0.997751380744095	FALSE
cb31828b8f23e5e0d3bec4436df5444c	111	0.997795403298972	FALSE
5323c6dd766d629cc750b0aabb4aaf72	92	0.997831444998761	FALSE
a3b72b215a67e3caf3e3db37446195a4	64	0.997949634878026	FALSE
138a326ebca05adbe3f688c7b3ae46b6	64	0.997949634878026	FALSE
09a077421a0f673d4a4ef0a0e3f7a913	64	0.997949634878026	FALSE
54f3fa8080e96ee8c48b77627ac0ae3a	114	0.998210893905748	FALSE
d77a3cc416991d5d3767214641e5b2ca	95	0.99823031082286	FALSE
2a5837baf8baa50071f3e36a0ff46fe2	236	0.99827897279397	FALSE
0a4009eb300a0f2671383587d6d8401e	115	0.998332246750527	FALSE
1815dac9ba55db9afc80283229a98eb4	115	0.998332246750527	FALSE
dbfde3e3bdbc75dd679758dad4834b73	115	0.998332246750527	FALSE
05b920e10cddf14fb5b0f0b42237d3e3	96	0.998347123119959	FALSE
57c580b06ef43b527dce14437166ba0b	136	0.99869184496278	FALSE
d226235298c561b8476ed3e44d6da337	100	0.998745655854517	FALSE
e9e8170bb8209aa94f8ec9a4718ffe6e	138	0.998870720871609	FALSE
698b4f5204ec0b08253a421ffd6bf57f	102	0.998909146884648	FALSE
fcb7d25194c7a7e800b5f7063a557582	102	0.998909146884648	FALSE
e3beabd067252bb387cfe12787744633	121	0.998913223732208	FALSE
bd2d39a2564b0ecc2f4e153d55d8fb7f	121	0.998913223732208	FALSE
4ad750cce3d83d64cabb192681982484	156	0.998990942515244	FALSE
366511529e7fcd4e72ba41ced9049d7b	140	0.999026645773671	FALSE
bec577c7cb4088fbb5631dc6ca656132	124	0.999126678027944	FALSE
76e476e3f5080af8597b9ea02dea6b0e	160	0.999258361933329	FALSE
5ebafbc67cbf3c6e47caca9b43ce08ef	108	0.999287599053711	FALSE
b5a4fee1a67d3818a8b8e8fbc9e19424	109	0.999337145076013	FALSE
e7ea4dc0277b1a7a2abe593bdc1d0ead	110	0.999383439652355	FALSE
c17057f15de467ed9c2c7a57817aa221	112	0.999467066146307	FALSE
9d63796fb581267f616c12eefa801ec2	115	0.999572774925443	FALSE
2068bc70de0f52d80b3aeaa003321213	151	0.999582373188258	FALSE
4ddd7df6ea0439b9b350e9795d3df725	134	0.999588342208228	FALSE
eb71a8778089c7404c228b317e38fef5	116	0.999603391690688	FALSE
ea587b40e7bc63eadc37e63ec3ac249e	152	0.999614271499125	FALSE
7e62ee3da1211add21aedb397c0807d7	184	0.999626238308179	FALSE
b2ea94f362ee35da22832eb8c68eb9dd	118	0.99965854857485	FALSE
acecad0c65e14937a0c7dbc6cb5f43c3	137	0.999673892453825	FALSE
c395ce96ac93ca3a6c81d07dbf85a94f	255	0.999684196681777	FALSE
e6ba559878b76dfc9c3c35aad453efb6	157	0.999742452486147	FALSE
e12ab73726e1064cf0896d411a38c0d4	140	0.999742581542467	FALSE
00ca181ebe6aba8dc69ca15436732300	140	0.999742581542467	FALSE
d9749f8b3a9efd705cf66b9e6c2a8e6f	140	0.999742581542467	FALSE
db8252bed6b71c5e2c72242fe02ce5b1	122	0.999747971981451	FALSE
cda4e6f933bb3108ea3e92f9db411c00	479	0.999771789839331	FALSE
bae67a3f8b760d01ac5d851c35de517e	142	0.999780578992592	FALSE
b36b52b10a3708b8e5fbaac814ab0931	272	0.999798373786883	FALSE
947a3c66fd46a8560854dd1c0dade2f3	126	0.999815044263348	FALSE
5343298f6ab7d59eb55c114cfec8b1ab	126	0.999815044263348	FALSE
5b4f8b625d8fbb1268863be7dbc4db5d	161	0.999815098254206	FALSE
3b555d2d7a26e295bdbdc64f10220157	147	0.999853885653837	FALSE
fe2f04ee87e40784f5104a15bec2058f	164	0.999856504232813	FALSE
60f2c9280ca556fd607cc51bedbbd8b6	148	0.99986547084978	FALSE
da77cbea962a956691c9df99fc91cccf	131	0.999875427733902	FALSE
61abd8a8fc64fc4ea6aa6bb33c056efd	166	0.999879113464545	FALSE
3cbaa88b2064dc5314799d9201534ec7	133	0.999893933187167	FALSE
fa60ad5919ebdc9b39b18872718a2393	290	0.99989482601865	FALSE
0d36c8f594b37214b4b58f29e1721f83	412	0.999899452336954	FALSE
c0e8688a1ac2878e2e6f62fe2a9304a0	134	0.999902186582871	FALSE
114a25e4cdc99c47c718434af3821ad4	135	0.999909834237843	FALSE
e59dec1a8a7c223eec18dbac3621bfcd	154	0.999918812324862	FALSE
a74d576891e4a308b4e08ab2a713bf5b	216	0.999922674573975	FALSE
7945babb9c10211ec6dc054af97441a7	190	0.999942767697508	FALSE
448df6a475edcccc5b389fe189f85fd5	147	0.999967158466949	FALSE
10036e06b250b7624c5b886014e3b2ec	147	0.999967158466949	FALSE
608e6548b1b4cbb6176c8fce090991a5	440	0.999969856753885	FALSE
b078843bd82f039e57f68ab3a55a39c1	265	0.99997139406525	FALSE
75dfd64719ae4652eac85de6f5bfe498	198	0.999973467632021	FALSE
eb01a38d86dcf9d0cc9d44abaa8152f4	213	0.999974597960593	FALSE
9cc12e6a4537bd55623ad90bf6b7084c	254	0.999976179830259	FALSE
ef606438f599842862bffdb564c7148d	154	0.99998231804312	FALSE
02bab2db0a7782cf48dec72cbac4bb20	172	0.999983924849906	FALSE
67a8bfa400cf64467ae73d78ee9b825c	188	0.999984067773002	FALSE
dfc9a1ff9b06e08d2132da15e61b3919	306	0.999985833424557	FALSE
0cfd5d23461f5a71d9628ac22ce590ac	295	0.99998649106477	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

5285ebad4f31b3e14f431007008acc63	235	0.999989412777277	FALSE
87cae47f30ea5a277a184645ac6c3b29	160	0.999989795870326	FALSE
d64d3900fff68de4ea004f18ba50c364	239	0.999993294835884	FALSE
a9a88c76960dfd5173b31ab28fe6828b	267	0.999995006944357	FALSE
29353ed4af9cbabb15cc3cc85a1f760a	215	0.999995510222954	FALSE
4c60dd7ba506649272fa67651b05a2f7	170	0.999996086406941	FALSE
b756bd4eff15dc5da0d88f9c34f65398	293	0.99999629171413	FALSE
d32274e6eeffddfd841ee35b5ba8ea1f	245	0.999996712798885	FALSE
ae8c85dfff97d56450e41e66c152ce85	271	0.999997027002998	FALSE
883d7985f03c15e9a5868bf9e14107d0	174	0.999997374135758	FALSE
babe77fccd08d91c140f37209e4bb5a5	220	0.999997444468929	FALSE
98679ce00beb048dfa099128eef70f21	175	0.99999762694986	FALSE
ad6f1612a5a7fdb80cf80a0402d48ee1	211	0.999998612465422	FALSE
b125cd368e6d851a7650f0a58795f5eb	181	0.99999872372356	FALSE
7ca9c8834010100a0dd569f0b2b5666b	213	0.999998898059349	FALSE
4cecd2478253cedd0dff65d35d7c434a	290	0.999998904915158	FALSE
a78fa572a60226d8150bfb0fc0fad652	241	0.999998942214219	FALSE
0a93c58cbaf5e764041cab0ad3e79b08	258	0.999999380746795	FALSE
50b47d356c39644171792954d17fce0b	188	0.999999398765722	FALSE
14fb9695d2e95bd663ad9a51b4d06723	276	0.999999726560132	FALSE
b78823261d78f2495896a69a3949e005	310	0.999999728244397	FALSE
773727ee7de31552dcf3f9609c18bdad	211	0.99999975236723	FALSE
28c1c628392df6a901829cfd2359a460	278	0.999999797358187	FALSE
f3a9e69bd296dd5d39ffecc544b36f25	398	0.99999988284644	FALSE
669fa3f5fbd7330211bc2d8d996a9e3e	220	0.999999919550495	FALSE
f5ef8bcbeee4753d8edc0b1d7e0aa499	206	0.999999926051858	FALSE
2081e8a45b69df88f6b3a5c914871247	240	0.999999965205849	FALSE
fb6931fb710455a0e1c19b26af270255	227	0.999999968102121	FALSE
5be8258fc43191236125bcc62765fdd9	293	0.999999982309437	FALSE
8cb3f4e315069c9334d1c069b96e3eac	234	0.999999987948591	FALSE
b8ba8f7a8bade720e87ca59fc4ab7fdb	274	0.999999991073874	FALSE
267d0f9a598522aa3ee4069277ea81bb	250	0.999999991988983	FALSE
6ace15b44dc69780c4bf45ef07ca096d	241	0.999999995678932	FALSE
b704e320a56fc1af1835929a3b71b7e3	255	0.999999996325366	FALSE
8114b1d0274e9e4bb6c91f6af1b8fac8	302	0.999999996590613	FALSE
3568d370fba8e070545cb84b03b7a088	268	0.999999996698289	FALSE
a2daf95d820d43e97bb2f8f61e3baaff	280	0.999999996796675	FALSE
de2a368a837144ad6e7bf64c176b742f	272	0.999999998330345	FALSE
fd75fae7f9721be5ba0e4a3704a1dedf	307	0.999999998724608	FALSE
ef54ef898867d22d2f65ebda83483ba1	331	0.999999999438227	FALSE
74299b73760426772762b92dde7d1131	322	0.999999999534007	FALSE
2eaaf1e130174d468d3b94cff4900e6e	281	0.99999999967167	FALSE
e7142feb387f929ee101779e66677cb9	297	0.999999999874447	FALSE
a2613697e6d6cc02bc530fd192f06bc9	311	0.999999999940768	FALSE
2f93e58b78f2842e83abd5fde37ad276	421	0.999999999980276	FALSE
2545c00c4b7992fda321d40d92a8cab4	336	0.99999999998447	FALSE
e083e2f58987c5f8db5d4dd16ddde91f	364	0.999999999990278	FALSE
5160485a724e51fcd3e4ebe658012aa8	416	0.999999999990541	FALSE
9e1d3dc63682fe2097c4dfa8dec76e9b	344	0.999999999999885	FALSE
24a60c6448e70d9198ad6ba93520958c	477	0.999999999999994	FALSE
28b8399c2b0790f21a60a17d495c620a	386	0.999999999999998	FALSE
520c77820886daeb8cf0d6497cfb1344	311	0.999999999999998	FALSE
5aa8281fbf9be4bfbf7ac852f1bdf8fe	373	0.999999999999999	FALSE
77accce4f7d181b78f7e66aad08abe01	439	1	FALSE
5f92443932d0deb4cf0d196b6e4fe4c0	396	1	FALSE
3f0d06b8c1b5e61b424af5b55ee3496d	411	1	FALSE
e92e9e21eab765c6ada40b76b7df1e0c	417	1	FALSE
523ac9f41cf23bf8c67d6537760b28d5	361	1	FALSE
49fc1baff761bf9d357844d43cd138ed	1	NA	FALSE
91d6bece60976a50b9051eea4bf40ad9	1	NA	FALSE
41916132a0320d7898614ab3368c5217	1	NA	FALSE
8811d5912a35b66ddd56a85f2f84756f	1	NA	FALSE
061cc3fd68a078dfc17c89dd4e7c1d57	1	NA	FALSE
f5b36d223324383df303174975ce46d5	1	NA	FALSE
36bb4e0630608c32e7b2a4b043c399c1	1	NA	FALSE
0984365862e8b81ae110762355cd5356	1	NA	FALSE
59f71f034d6f5f3211ce16437c08a399	1	NA	FALSE
8dcdc18ee9b372e12ad9acf3b560f0de	1	NA	FALSE
f42582c91ba40de887b64d7b2356d6d8	1	NA	FALSE
dbd6a52a5c2d8dd052e383c067109d73	1	NA	FALSE
3540e1ab31453e650aed7b1c16c68373	1	NA	FALSE
ae37473e32867acb33cd941ca2a0bc9f	1	NA	FALSE
6639175de58b5f0ee31f8300c07adc72	1	NA	FALSE
8cf34b55c0736c2a6e94f72fd0aa07e1	1	NA	FALSE
408f662e9702a6252066a482053dfa05	1	NA	FALSE
0556309e565282e3cdc6d22c7a1314ca	1	NA	FALSE
a31397646d05a84f8890b670601c2a88	1	NA	FALSE
c6e6c1884afc6be4ad05add5339cc95e	1	NA	FALSE
1e3ca86e4fa19f2b96529f8a78b6304a	1	NA	FALSE
90a7ad77ba874fb1ad37e15e4defbb91	1	NA	FALSE
6e0436014e7f545bb0c89c6f90bf2c82	1	NA	FALSE
1488392be87091b8779262c95753d489	1	NA	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

71a639b66d4d99791d5c69e649932498	1	NA	FALSE
b90a3f4af43dd1eadaac36033d825b32	1	NA	FALSE
7c085dcaa22450a60d0bcdcff9c02f1d	1	NA	FALSE
7615b39e73064bf612abefe1336a5f5e	1	NA	FALSE
0cf91d09851e5165515a95c48f990b2c	1	NA	FALSE
40ef360e3de9eef2e54cc384c0f442fb	1	NA	FALSE
c9d2e62830d879d1b703ec112804ed4f	1	NA	FALSE
179dfc0f76a5c60ed0db3f36ea4103d5	1	NA	FALSE
59bbb81117691ae8817ac815f73b9945	1	NA	FALSE
27cee38f3d317da09942584a9e611701	1	NA	FALSE
a344d280d5637fa174318df87463e8b2	1	NA	FALSE
1abf91747a161b143a6733473add3d1f	1	NA	FALSE
58172f93459801cf780471559743a1a1	1	NA	FALSE
2dbe816b5bd581af1d9e40b426a3ccc5	1	NA	FALSE
ad213ddb834c49ba5c3fb2c35715ea31	1	NA	FALSE
f4759806bae7e262721beb08f96af6f2	1	NA	FALSE
9f6cc04b4b4b011455a72b5a60460c23	1	NA	FALSE
c3a7e5679db8cc9a7dcb122d21a62abd	1	NA	FALSE
99655509650822e576d12e2203109953	1	NA	FALSE
c4a5fa167be7022d7e8f6a24ea5a814e	1	NA	FALSE
a75568deb408dcc5be74bdf23e1f6dc7	1	NA	FALSE
6ac5bf8ed56c7709a77fc026615798ad	1	NA	FALSE
324760267b2fe263a0c835e9af329c00	1	NA	FALSE
6f5e0038f0b835752a599ff8ec9801ea	1	NA	FALSE
88d21bfdf6bb7dd539b1f97ef3fd3d11	1	NA	FALSE
7148747d2d38f0d895386df0c931dcb4	1	NA	FALSE
fa50250efdc5d41bd6577db321412f03	1	NA	FALSE
8d97a06ddc80deb17af3ed88f954f8e1	1	NA	FALSE
f68c7cd7566c0f1295c0e531c09b5c1f	1	NA	FALSE
8c3070a01c5d1017b27e2dac97270d05	1	NA	FALSE
7425d153990b97e5f0e131d34cad5aa4	1	NA	FALSE
c1d232ebefcae769cfd656b0931cd11a	1	NA	FALSE
7591056a10b3bd4cc9621543330349ef	1	NA	FALSE
20c1ba0382813ccc891453a727f45d2d	1	NA	FALSE
ced6caa0b6c0ffc98a8414b5bdc83112	1	NA	FALSE
3f1c0b87de5bb4df2b8d2e6b36bed0ab	1	NA	FALSE
f6799ce3e138384ebd9119e29709a3f3	1	NA	FALSE
855b88b34ea95ec7b5c3de4ad4c824ee	1	NA	FALSE
30adb700aa2336a724166e014d162e7f	1	NA	FALSE
5cd044dfb5feca3ee72dae76fa6e9ee9	1	NA	FALSE
5e6f632ca87cf43563b74e15539e7785	1	NA	FALSE
bd96e429d73d57cbb9c8ef9bc46b7820	1	NA	FALSE
65e2fe387bfdf0fa880811cceac1d59e	1	NA	FALSE
5a14b8ff1aecdf0489c6820e8e105b88	1	NA	FALSE
ef5a2480e55d717d0c565e287f82a473	1	NA	FALSE
9ae9c5171312ae7ce382a7baa6b81362	1	NA	FALSE
0911bb821081272792779660 66a861651	1	NA	FALSE
85a7a27464dba46cd60d893ad0b3d35f	1	NA	FALSE
3114f9c7de279de58fe3dc15e2dd14fb	1	NA	FALSE
b4bee6fb68e88d00ba80c1bc627510c0	1	NA	FALSE
12679b50e6d1202949122b8b0622ade1	1	NA	FALSE
9f2cf0e7035d0f777c21936203005b14	1	NA	FALSE
394c2ec9cae1af76e0b585fedacb13a2	1	NA	FALSE
8bedbd900998bc8be0f4e32027c14b79	1	NA	FALSE
d91b9586180e2dc67db8d50bfc06d00c	1	NA	FALSE
9ae3e5447f15b13586d0d3803b8e68af	1	NA	FALSE
e50354b9c4d97287a20c1e167dfefa3e	1	NA	FALSE
ba76aae4914cd1c68039931188d050fe	1	NA	FALSE
03b679010b1bf8b116785f42667e02e1	1	NA	FALSE
e7eb03a5b2a368b8a6099d242bcacd54	1	NA	FALSE
945f64283b2eba0afa876de55b78ef31	1	NA	FALSE
c686148b8d01f84f7604bc5fcaddbebd	1	NA	FALSE
0407fd8be2c9084dafdf0c2cd3d0befa	1	NA	FALSE
8d2de3aea083cf4f7d1eae80c2af9146	1	NA	FALSE
3e23230e4686b96576316f5e033b7082	1	NA	FALSE
6e7e1fbac8c5e839a9cdfc5fba8f7770	1	NA	FALSE
7b3fa6d9e1062f0b6537b7372012b859	1	NA	FALSE
d19aea26e0ad7c2155fdd287a2ec6b77	1	NA	FALSE
0ca49b13050f305492d115000a1d001d	1	NA	FALSE
2b9cede5145c126ca5a5de9fc895a14b	1	NA	FALSE
1a902f04710149ed9498c3e99a5afd63	1	NA	FALSE
8577af56ac1b94a1aa8b74077b5a5cbc	1	NA	FALSE
bd1a351302ec362f46db2979bada621d	1	NA	FALSE
352144a7596df7857029b20703ae2eda	1	NA	FALSE
5c3c19675e638054dd3325e5ba5677d2	1	NA	FALSE
15d139a08d6c1648ba4c20c9995ab9b2	1	NA	FALSE
2ca106d719b5911c71a6e189459efa66	1	NA	FALSE
845f586ce16f4bd61ab41379809ae1e7	1	NA	FALSE
72b07965cf4f08eff549d0053d0906f1	1	NA	FALSE
1b336d762092361e9f7c04a56548a6a2	1	NA	FALSE
2429b0edc950114d01ddec0d631166cd	1	NA	FALSE
3f839cccd6e9568fcb569b0e4f7d6ee8	1	NA	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

1c0e84ab16636a5d983a7f210354520b	1	NA	FALSE
f7163bb266bce765f0aa145cf7a6c37f	1	NA	FALSE
6fd9843eb698930e30818d77d505ed5c	1	NA	FALSE
94b612176f3b4cea22cda2f67134ebe6	1	NA	FALSE
25ddb3b811c0140bec907e45cf4b2a45	1	NA	FALSE
68f55b94cf23a2b444b7c8929907eb7f	1	NA	FALSE
97f006e854e73edfcecc1415ac23de6d	1	NA	FALSE
8171a8faf564215283a3fccfe23a4026	1	NA	FALSE
3eac2f1f0d1c5a98e4dcfe3bb7c16709	1	NA	FALSE
832fadfb5fe2b9bec572b5788cfb3f73	1	NA	FALSE
7a538dcb40656b6be4c267503c501b0a	1	NA	FALSE
4e4b6a3260a6061be4590ed7fd29a214	1	NA	FALSE
5ed7d64c93da6dfb32f43400d14ca8e3	1	NA	FALSE
3e92de515639a37692e35e261af54e98	1	NA	FALSE
1848ba16ac6e038d4be98fd5d7bfead2	1	NA	FALSE
bb5e5ae0efa9bf009dd646ebcac3bab4	1	NA	FALSE
703968088ed15033df6c62f2530e50b3	1	NA	FALSE
62107a3ffb0f00f6ff881660f970e449	1	NA	FALSE
d3f91dd2f8e03e272362cb77276b159c	1	NA	FALSE
37e1ba9d0783625524befd620d1e3d29	1	NA	FALSE
3d6c6053fcc0b58354ac9dbe2d2f699e	1	NA	FALSE

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

	#16s_sampleID		Subject	Monthcat	Group	mBorg	6MWT	CAT
Sample_1	16s_sampl	3R451_FU	3R451_FU	subject3R4	M0	Control	Not apply	Not apply	Not apply
Sample_2	16s_sampl	3R452_FU	3R452_FU	subject3R4	M0	Control	Not apply	Not apply	Not apply
Sample_3	16s_sampl	3R_455_M	3R_455_M	subject3R4	M0	Control	Not apply	Not apply	Not apply
Sample_4	16s_sampl	508_FU1_	508_FU1_	subject508	M0	Control	Not apply	Not apply	Not apply
Sample_5	16s_sampl	509_FU1_	509_FU1_	subject509	M0	Control	Not apply	Not apply	Not apply
Sample_6	16s_sampl	519_M1_p	519_M1_p	subject519	M0	Control	Not apply	Not apply	Not apply
Sample_7	16s_sampl	526_FU1_	526_FU1_	subject526	M0	Control	Not apply	Not apply	Not apply
Sample_8	16s_sampl	527_FU1_	527_FU1_	subject527	M0	Control	Not apply	Not apply	Not apply
Sample_9	16s_sampl	540_M1_L	540_M1_L	subject540	M0	Control	Not apply	Not apply	Not apply
Sample_10	16s_sampl	578_M1_L	578_M1_L	subject578	M0	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	583_FU1_	583_FU1_	subject583	M0	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	649_FU1_	649_FU1_	subject649	M0	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	663_M1_L	663_M1_L	subject663	M0	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	665_M1_p	665_M1_p	subject665	M0	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	676_M1_L	676_M1_L	subject676	M0	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	678_M1_p	678_M1_p	subject678	M0	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	694_M1_p	694_M1_p	subject694	M0	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	695_M1_L	695_M1_L	subject695	M0	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	794_FU1_	794_FU1_	subject794	M0	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	795_FU1_	795_FU1_	subject795	M0	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	801_FU1_	801_FU1_	subject801	M0	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	807_FU1_	807_FU1_	subject807	M0	Control	Not apply	Not apply	Not apply
Sample_23	16s_sampl	809_FU1_	809_FU1_	subject809	M0	Control	Not apply	Not apply	Not apply
Sample_24	16s_sampl	830_FU1_	830_FU1_	subject830	M0	Control	Not apply	Not apply	Not apply
Sample_25	16s_sampl	833_FU1_	833_FU1_	subject833	M0	Control	Not apply	Not apply	Not apply
Sample_26	16s_sampl	837_M1_L	837_M1_L	subject837	M0	Control	Not apply	Not apply	Not apply
Sample_27	16s_sampl	851_FU1_	851_FU1_	subject851	M0	Control	Not apply	Not apply	Not apply
Sample_28	16s_sampl	880_M1_L	880_M1_L	subject880	M0	Control	Not apply	Not apply	Not apply
Sample_29	16s_sampl	884_M1_L	884_M1_L	subject884	M0	Control	Not apply	Not apply	Not apply
Sample_30	16s_sampl	885_M1_p	885_M1_p	subject885	M0	Control	Not apply	Not apply	Not apply
Sample_31	16s_sampl	891_M1_L	891_M1_L	subject891	M0	Control	Not apply	Not apply	Not apply
Sample_32	16s_sampl	897_M1_L	897_M1_L	subject897	M0	Control	Not apply	Not apply	Not apply
Sample_33	16s_sampl	907_M1_L	907_M1_L	subject907	M0	Control	Not apply	Not apply	Not apply
Sample_34	16s_sampl	934_M1_L	934_M1_L	subject934	M0	Control	Not apply	Not apply	Not apply
Sample_35	16s_sampl	935_FU1_	935_FU1_	subject935	M0	Control	Not apply	Not apply	Not apply
Sample_36	16s_sampl	P101_FU1_	P101_FU1_	subjectP10	M0	Control	Not apply	Not apply	Not apply
Sample_37	16s_sampl	P201_FU1_	P201_FU1_	subjectP20	M0	Control	Not apply	Not apply	Not apply
Sample_38	16s_sampl	P204_FU1_	P204_FU1_	subjectP20	M0	Control	Not apply	Not apply	Not apply
Sample_39	16s_sampl	3R451_FU	3R451_FU	subject3R4	M1	Control	Not apply	Not apply	Not apply
Sample_40	16s_sampl	3R452_FU	3R452_FU	subject3R4	M1	Control	Not apply	Not apply	Not apply
Sample_41	16s_sampl	3R_455_M	3R_455_M	subject3R4	M1	Control	Not apply	Not apply	Not apply
Sample_42	16s_sampl	508_FU2_	508_FU2_	subject508	M1	Control	Not apply	Not apply	Not apply
Sample_43	16s_sampl	509_FU2_	509_FU2_	subject509	M1	Control	Not apply	Not apply	Not apply
Sample_44	16s_sampl	519_M2_p	519_M2_p	subject519	M1	Control	Not apply	Not apply	Not apply
Sample_45	16s_sampl	526_FU2_	526_FU2_	subject526	M1	Control	Not apply	Not apply	Not apply
Sample_46	16s_sampl	527_FU2_	527_FU2_	subject527	M1	Control	Not apply	Not apply	Not apply
Sample_47	16s_sampl	540_M2_L	540_M2_L	subject540	M1	Control	Not apply	Not apply	Not apply
Sample_48	16s_sampl	578_M2_L	578_M2_L	subject578	M1	Control	Not apply	Not apply	Not apply
Sample_49	16s_sampl	583_FU2_	583_FU2_	subject583	M1	Control	Not apply	Not apply	Not apply
Sample_50	16s_sampl	649_FU2_	649_FU2_	subject649	M1	Control	Not apply	Not apply	Not apply

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_51	16s_sampl	663_M2_L	663_M2_L	subject663	M1	Control	Not apply	Not apply	Not apply
Sample_52	16s_sampl	665_M2_p	665_M2_p	subject665	M1	Control	Not apply	Not apply	Not apply
Sample_53	16s_sampl	676_M2_L	676_M2_L	subject676	M1	Control	Not apply	Not apply	Not apply
Sample_54	16s_sampl	678_M2_p	678_M2_p	subject678	M1	Control	Not apply	Not apply	Not apply
Sample_55	16s_sampl	694_M2_p	694_M2_p	subject694	M1	Control	Not apply	Not apply	Not apply
Sample_56	16s_sampl	695_M2_L	695_M2_L	subject695	M1	Control	Not apply	Not apply	Not apply
Sample_57	16s_sampl	794_FU2_	794_FU2_	subject794	M1	Control	Not apply	Not apply	Not apply
Sample_58	16s_sampl	795_FU2_	795_FU2_	subject795	M1	Control	Not apply	Not apply	Not apply
Sample_59	16s_sampl	801_FU2_	801_FU2_	subject801	M1	Control	Not apply	Not apply	Not apply
Sample_60	16s_sampl	807_FU2_	807_FU2_	subject807	M1	Control	Not apply	Not apply	Not apply
Sample_61	16s_sampl	809_FU2_	809_FU2_	subject809	M1	Control	Not apply	Not apply	Not apply
Sample_62	16s_sampl	830_FU2_	830_FU2_	subject830	M1	Control	Not apply	Not apply	Not apply
Sample_63	16s_sampl	833_FU2_	833_FU2_	subject833	M1	Control	Not apply	Not apply	Not apply
Sample_64	16s_sampl	837_M2_L	837_M2_L	subject837	M1	Control	Not apply	Not apply	Not apply
Sample_65	16s_sampl	851_FU2_	851_FU2_	subject851	M1	Control	Not apply	Not apply	Not apply
Sample_66	16s_sampl	880_M2_L	880_M2_L	subject880	M1	Control	Not apply	Not apply	Not apply
Sample_67	16s_sampl	884_M2_L	884_M2_L	subject884	M1	Control	Not apply	Not apply	Not apply
Sample_68	16s_sampl	885_M2_p	885_M2_p	subject885	M1	Control	Not apply	Not apply	Not apply
Sample_69	16s_sampl	891_M2_L	891_M2_L	subject891	M1	Control	Not apply	Not apply	Not apply
Sample_70	16s_sampl	897_M2_L	897_M2_L	subject897	M1	Control	Not apply	Not apply	Not apply
Sample_71	16s_sampl	907_M2_L	907_M2_L	subject907	M1	Control	Not apply	Not apply	Not apply
Sample_72	16s_sampl	934_M2_L	934_M2_L	subject934	M1	Control	Not apply	Not apply	Not apply
Sample_73	16s_sampl	935_FU2_	935_FU2_	subject935	M1	Control	Not apply	Not apply	Not apply
Sample_74	16s_sampl	P101_FU2_	P101_FU2_	subjectP10	M1	Control	Not apply	Not apply	Not apply
Sample_75	16s_sampl	P201_FU2_	P201_FU2_	subjectP20	M1	Control	Not apply	Not apply	Not apply
Sample_76	16s_sampl	P204_FU2_	P204_FU2_	subjectP20	M1	Control	Not apply	Not apply	Not apply
Sample_77	16s_sampl	3R451_FU	3R451_FU	subject3R4	M2	Control	Not apply	Not apply	Not apply
Sample_78	16s_sampl	3R452_FU	3R452_FU	subject3R4	M2	Control	Not apply	Not apply	Not apply
Sample_79	16s_sampl	3R_455_M	3R_455_M	subject3R4	M2	Control	Not apply	Not apply	Not apply
Sample_80	16s_sampl	508_FU3_	508_FU3_	subject508	M2	Control	Not apply	Not apply	Not apply
Sample_81	16s_sampl	509_FU3_	509_FU3_	subject509	M2	Control	Not apply	Not apply	Not apply
Sample_82	16s_sampl	519_M3_p	519_M3_p	subject519	M2	Control	Not apply	Not apply	Not apply
Sample_83	16s_sampl	526_FU3_	526_FU3_	subject526	M2	Control	Not apply	Not apply	Not apply
Sample_84	16s_sampl	527_FU3_	527_FU3_	subject527	M2	Control	Not apply	Not apply	Not apply
Sample_85	16s_sampl	540_M3_L	540_M3_L	subject540	M2	Control	Not apply	Not apply	Not apply
Sample_86	16s_sampl	578_M3_L	578_M3_L	subject578	M2	Control	Not apply	Not apply	Not apply
Sample_87	16s_sampl	583_FU3_	583_FU3_	subject583	M2	Control	Not apply	Not apply	Not apply
Sample_88	16s_sampl	649_FU3_	649_FU3_	subject649	M2	Control	Not apply	Not apply	Not apply
Sample_89	16s_sampl	663_M3_L	663_M3_L	subject663	M2	Control	Not apply	Not apply	Not apply
Sample_90	16s_sampl	665_M3_p	665_M3_p	subject665	M2	Control	Not apply	Not apply	Not apply
Sample_91	16s_sampl	676_M3_L	676_M3_L	subject676	M2	Control	Not apply	Not apply	Not apply
Sample_92	16s_sampl	678_M3_p	678_M3_p	subject678	M2	Control	Not apply	Not apply	Not apply
Sample_93	16s_sampl	694_M3_p	694_M3_p	subject694	M2	Control	Not apply	Not apply	Not apply
Sample_94	16s_sampl	695_M3_L	695_M3_L	subject695	M2	Control	Not apply	Not apply	Not apply
Sample_95	16s_sampl	794_FU3_	794_FU3_	subject794	M2	Control	Not apply	Not apply	Not apply
Sample_96	16s_sampl	795_FU3_	795_FU3_	subject795	M2	Control	Not apply	Not apply	Not apply
Sample_97	16s_sampl	801_FU3_	801_FU3_	subject801	M2	Control	Not apply	Not apply	Not apply
Sample_98	16s_sampl	807_FU3_	807_FU3_	subject807	M2	Control	Not apply	Not apply	Not apply
Sample_99	16s_sampl	809_FU3_	809_FU3_	subject809	M2	Control	Not apply	Not apply	Not apply
Sample_10	16s_sampl	830_FU3_	830_FU3_	subject830	M2	Control	Not apply	Not apply	Not apply
Sample_10	16s_sampl	833_FU3_	833_FU3_	subject833	M2	Control	Not apply	Not apply	Not apply

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_1C	16s_sampl	837_M3_L	837_M3_L	subject837	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	851_FU3_p	851_FU3_p	subject851	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	880_M3_L	880_M3_L	subject880	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	884_M3_L	884_M3_L	subject884	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	885_M3_p	885_M3_p	subject885	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	891_M3_L	891_M3_L	subject891	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	897_M3_L	897_M3_L	subject897	M2	Control	Not apply	Not apply	Not apply
Sample_1C	16s_sampl	907_M3_L	907_M3_L	subject907	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	934_M3_L	934_M3_L	subject934	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	935_FU3_p	935_FU3_p	subject935	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	P101_FU3_	P101_FU3_	subjectP10	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	P201_FU3_	P201_FU3_	subjectP20	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	P204_FU3_	P204_FU3_	subjectP20	M2	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	3R451_FU4	3R451_FU4	subject3R4	M3	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	3R452_FU4	3R452_FU4	subject3R4	M3	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	3R_455_M	3R_455_M	subject3R4	M3	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	508_FU4_p	508_FU4_p	subject508	M3	Control	Not apply	Not apply	Not apply
Sample_11	16s_sampl	509_FU4_p	509_FU4_p	subject509	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	519_M4_p	519_M4_p	subject519	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	526_FU4_p	526_FU4_p	subject526	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	527_FU4_p	527_FU4_p	subject527	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	540_M4_L	540_M4_L	subject540	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	578_M4_L	578_M4_L	subject578	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	583_FU4_p	583_FU4_p	subject583	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	649_FU4_p	649_FU4_p	subject649	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	663_M4_L	663_M4_L	subject663	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	665_M4_p	665_M4_p	subject665	M3	Control	Not apply	Not apply	Not apply
Sample_12	16s_sampl	676_M4_L	676_M4_L	subject676	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	678_M4_p	678_M4_p	subject678	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	694_M4_p	694_M4_p	subject694	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	695_M4_L	695_M4_L	subject695	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	794_FU4_p	794_FU4_p	subject794	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	795_FU4_p	795_FU4_p	subject795	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	801_FU4_p	801_FU4_p	subject801	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	807_FU4_p	807_FU4_p	subject807	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	809_FU4_p	809_FU4_p	subject809	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	830_FU4_p	830_FU4_p	subject830	M3	Control	Not apply	Not apply	Not apply
Sample_13	16s_sampl	833_FU4_p	833_FU4_p	subject833	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	837_M4_L	837_M4_L	subject837	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	851_FU4_p	851_FU4_p	subject851	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	880_M4_L	880_M4_L	subject880	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	884_M4_L	884_M4_L	subject884	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	885_M4_p	885_M4_p	subject885	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	891_M4_L	891_M4_L	subject891	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	897_M4_L	897_M4_L	subject897	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	907_M4_L	907_M4_L	subject907	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	934_M4_L	934_M4_L	subject934	M3	Control	Not apply	Not apply	Not apply
Sample_14	16s_sampl	935_FU4_p	935_FU4_p	subject935	M3	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	P101_FU4_	P101_FU4_	subjectP10	M3	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	P201_FU4_	P201_FU4_	subjectP20	M3	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	P204_FU4_	P204_FU4_	subjectP20	M3	Control	Not apply	Not apply	Not apply

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_15	16s_sampl	3R451_FU!	3R451_FU!	subject3R4	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	3R452_FU!	3R452_FU!	subject3R4	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	3R_455_M	3R_455_M	subject3R4	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	508_FU5_p	508_FU5_p	subject508	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	509_FU5_p	509_FU5_p	subject509	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	519_M5_p	519_M5_p	subject519	M4	Control	Not apply	Not apply	Not apply
Sample_15	16s_sampl	526_FU5_p	526_FU5_p	subject526	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	527_FU5_p	527_FU5_p	subject527	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	540_M5_L	540_M5_L	subject540	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	578_M5_L	578_M5_L	subject578	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	583_FU5_p	583_FU5_p	subject583	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	649_FU5_p	649_FU5_p	subject649	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	663_M5_L	663_M5_L	subject663	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	665_M5_p	665_M5_p	subject665	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	676_M5_L	676_M5_L	subject676	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	678_M5_p	678_M5_p	subject678	M4	Control	Not apply	Not apply	Not apply
Sample_16	16s_sampl	694_M5_p	694_M5_p	subject694	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	695_M5_L	695_M5_L	subject695	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	794_FU5_p	794_FU5_p	subject794	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	795_FU5_p	795_FU5_p	subject795	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	801_FU5_p	801_FU5_p	subject801	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	807_FU5_p	807_FU5_p	subject807	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	809_FU5_p	809_FU5_p	subject809	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	830_FU5_p	830_FU5_p	subject830	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	833_FU5_p	833_FU5_p	subject833	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	837_M5_L	837_M5_L	subject837	M4	Control	Not apply	Not apply	Not apply
Sample_17	16s_sampl	851_FU5_p	851_FU5_p	subject851	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	880_M5_L	880_M5_L	subject880	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	884_M5_L	884_M5_L	subject884	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	885_M5_p	885_M5_p	subject885	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	891_M5_L	891_M5_L	subject891	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	897_M5_L	897_M5_L	subject897	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	907_M5_L	907_M5_L	subject907	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	934_M5_L	934_M5_L	subject934	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	935_FU5_p	935_FU5_p	subject935	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	P101_FU5_	P101_FU5_	subjectP10	M4	Control	Not apply	Not apply	Not apply
Sample_18	16s_sampl	P201_FU5_	P201_FU5_	subjectP20	M4	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	P204_FU5_	P204_FU5_	subjectP20	M4	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	3R451_FU(	3R451_FU(	subject3R4	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	3R452_FU(	3R452_FU(	subject3R4	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	3R_455_M	3R_455_M	subject3R4	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	508_FU6_p	508_FU6_p	subject508	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	509_FU6_p	509_FU6_p	subject509	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	519_M6_p	519_M6_p	subject519	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	526_FU6_p	526_FU6_p	subject526	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	527_FU6_p	527_FU6_p	subject527	M5	Control	Not apply	Not apply	Not apply
Sample_19	16s_sampl	540_M6_L	540_M6_L	subject540	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	578_M6_L	578_M6_L	subject578	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	583_FU6_p	583_FU6_p	subject583	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	649_FU6_p	649_FU6_p	subject649	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	663_M6_L	663_M6_L	subject663	M5	Control	Not apply	Not apply	Not apply

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_20	16s_sampl	665_M6_p	665_M6_p	subject665	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	676_M6_L	676_M6_L	subject676	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	678_M6_p	678_M6_p	subject678	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	694_M6_p	694_M6_p	subject694	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	695_M6_L	695_M6_L	subject695	M5	Control	Not apply	Not apply	Not apply
Sample_20	16s_sampl	794_FU6_p	794_FU6_p	subject794	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	795_FU6_p	795_FU6_p	subject795	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	801_FU6_p	801_FU6_p	subject801	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	807_FU6_p	807_FU6_p	subject807	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	809_FU6_p	809_FU6_p	subject809	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	830_FU6_p	830_FU6_p	subject830	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	833_FU6_p	833_FU6_p	subject833	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	837_M6_L	837_M6_L	subject837	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	851_FU6_p	851_FU6_p	subject851	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	880_M6_L	880_M6_L	subject880	M5	Control	Not apply	Not apply	Not apply
Sample_21	16s_sampl	884_M6_L	884_M6_L	subject884	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	885_M6_p	885_M6_p	subject885	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	891_M6_L	891_M6_L	subject891	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	897_M6_L	897_M6_L	subject897	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	907_M6_L	907_M6_L	subject907	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	934_M6_L	934_M6_L	subject934	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	935_FU6_p	935_FU6_p	subject935	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	P101_FU6_	P101_FU6_	subjectP10	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	P201_FU6_	P201_FU6_	subjectP20	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	P204_FU6_	P204_FU6_	subjectP20	M5	Control	Not apply	Not apply	Not apply
Sample_22	16s_sampl	3R443_M3	3R443_M3	subject443	M0	interventic	NR	R	NR
Sample_23	16s_sampl	3R474_M3	3R474_M3	subject474	M0	interventic	NR	R	NR
Sample_23	16s_sampl	3R478_M3	3R478_M3	subject478	M0	interventic	NR	NR	R
Sample_23	16s_sampl	3R479_M3	3R479_M3	subject479	M0	interventic	NR	NR	R
Sample_23	16s_sampl	3R481_M3	3R481_M3	subject481	M0	interventic	NR	NR	R
Sample_23	16s_sampl	3R482_M3	3R482_M3	subject482	M0	interventic	NR	NR	NR
Sample_23	16s_sampl	3R486_M3	3R486_M3	subject486	M0	interventic	NR	R	R
Sample_23	16s_sampl	3R487_M3	3R487_M3	subject487	M0	interventic	R	R	R
Sample_23	16s_sampl	3R488_M3	3R488_M3	subject488	M0	interventic	R	R	NR
Sample_23	16s_sampl	3R492_M3	3R492_M3	subject492	M0	interventic	NR	R	NR
Sample_23	16s_sampl	3R493_M3	3R493_M3	subject493	M0	interventic	NR	R	R
Sample_24	16s_sampl	3R495_M3	3R495_M3	subject495	M0	interventic	NR	R	NR
Sample_24	16s_sampl	3R496_M3	3R496_M3	subject496	M0	interventic	NR	NR	NR
Sample_24	16s_sampl	3R506_M3	3R506_M3	subject506	M0	interventic	NR	NR	NR
Sample_24	16s_sampl	3R512_M3	3R512_M3	subject512	M0	interventic	NR	NR	R
Sample_24	16s_sampl	863_M1_L	863_M1_L	subject863	M0	interventic	NR	R	R
Sample_24	16s_sampl	P10_M1_p	P10_M1_p	subject10	M0	interventic	R	R	R
Sample_24	16s_sampl	P115_M1_	P115_M1_	subjectP11	M0	interventic	NR	R	R
Sample_24	16s_sampl	P11_M3_p	P11_M3_p	subject11	M0	interventic	NR	NR	R
Sample_24	16s_sampl	P121_M1_	P121_M1_	subjectP12	M0	interventic	NR	R	NR
Sample_24	16s_sampl	P125_M1_	P125_M1_	subjectP12	M0	interventic	NR	R	R
Sample_25	16s_sampl	P127_M1_	P127_M1_	subjectP12	M0	interventic	NR	R	R
Sample_25	16s_sampl	P129_M1_	P129_M1_	subjectP12	M0	interventic	NR	R	R
Sample_25	16s_sampl	P13_M3_p	P13_M3_p	subjectP13	M0	interventic	NR	NR	R
Sample_25	16s_sampl	P13_M7_p	P13_M7_p	subject13	M0	interventic	NR	R	R
Sample_25	16s_sampl	P140_M1_	P140_M1_	subjectP14	M0	interventic	NR	NR	NR

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_25	16s_sampl	P143_M1_	P143_M1_	subjectP14	M0	interventic	R	R	R
Sample_25	16s_sampl	P145_M1_	P145_M1_	subjectP14	M0	interventic	R	R	R
Sample_25	16s_sampl	P147_M1_	P147_M1_	subjectP14	M0	interventic	R	R	NR
Sample_25	16s_sampl	P149_M1_	P149_M1_	subjectP14	M0	interventic	NR	R	R
Sample_25	16s_sampl	P150_M1_	P150_M1_	subjectP15	M0	interventic	R	NR	NR
Sample_26	16s_sampl	P16_M3_p	P16_M3_p	subject16	M0	interventic	NR	R	R
Sample_26	16s_sampl	P18_M3_p	P18_M3_p	subject18	M0	interventic	NR	NR	NR
Sample_26	16s_sampl	P19_M3_p	P19_M3_p	subject19	M0	interventic	NR	NR	R
Sample_26	16s_sampl	P1_M3_ph	P1_M3_ph	subject1	M0	interventic	NR	NR	NR
Sample_26	16s_sampl	P2_M3_ph	P2_M3_ph	subject2	M0	interventic	R	R	R
Sample_26	16s_sampl	P6_M3_ph	P6_M3_ph	subject6	M0	interventic	NR	R	R
Sample_26	16s_sampl	P7_M3_ph	P7_M3_ph	subject7	M0	interventic	R	R	R
Sample_26	16s_sampl	3R443_M4	3R443_M4	subject443	M1	interventic	NR	R	NR
Sample_26	16s_sampl	3R474_M4	3R474_M4	subject474	M1	interventic	NR	R	NR
Sample_26	16s_sampl	3R478_M4	3R478_M4	subject478	M1	interventic	NR	NR	R
Sample_27	16s_sampl	3R481_M4	3R481_M4	subject481	M1	interventic	NR	NR	R
Sample_27	16s_sampl	3R482_M4	3R482_M4	subject482	M1	interventic	NR	NR	NR
Sample_27	16s_sampl	3R486_M4	3R486_M4	subject486	M1	interventic	NR	R	R
Sample_27	16s_sampl	3R487_M4	3R487_M4	subject487	M1	interventic	R	R	R
Sample_27	16s_sampl	3R488_M4	3R488_M4	subject488	M1	interventic	R	R	NR
Sample_27	16s_sampl	3R492_M4	3R492_M4	subject492	M1	interventic	NR	R	NR
Sample_27	16s_sampl	3R493_M4	3R493_M4	subject493	M1	interventic	NR	R	R
Sample_27	16s_sampl	3R495_M4	3R495_M4	subject495	M1	interventic	NR	R	NR
Sample_27	16s_sampl	3R496_M4	3R496_M4	subject496	M1	interventic	NR	NR	NR
Sample_27	16s_sampl	3R506_M4	3R506_M4	subject506	M1	interventic	NR	NR	NR
Sample_28	16s_sampl	3R512_M4	3R512_M4	subject512	M1	interventic	NR	NR	R
Sample_28	16s_sampl	863_M3_L	863_M3_L	subject863	M1	interventic	NR	R	R
Sample_28	16s_sampl	P10_M4_p	P10_M4_p	subject10	M1	interventic	R	R	R
Sample_28	16s_sampl	P115_M3_	P115_M3_	subjectP11	M1	interventic	NR	R	R
Sample_28	16s_sampl	P11_M4_p	P11_M4_p	subject11	M1	interventic	NR	NR	R
Sample_28	16s_sampl	P121_M3_	P121_M3_	subjectP12	M1	interventic	NR	R	NR
Sample_28	16s_sampl	P125_M3_	P125_M3_	subjectP12	M1	interventic	NR	R	R
Sample_28	16s_sampl	P127_M3_	P127_M3_	subjectP12	M1	interventic	NR	R	R
Sample_28	16s_sampl	P129_M3_	P129_M3_	subjectP12	M1	interventic	NR	R	R
Sample_28	16s_sampl	P13_M5_p	P13_M5_p	subjectP13	M1	interventic	NR	NR	R
Sample_29	16s_sampl	P134_M1_	P134_M1_	subject13	M1	interventic	NR	R	R
Sample_29	16s_sampl	P140_M3_	P140_M3_	subjectP14	M1	interventic	NR	NR	NR
Sample_29	16s_sampl	P143_M3_	P143_M3_	subjectP14	M1	interventic	R	R	R
Sample_29	16s_sampl	P145_M3_	P145_M3_	subjectP14	M1	interventic	R	R	R
Sample_29	16s_sampl	P147_M3_	P147_M3_	subjectP14	M1	interventic	R	R	NR
Sample_29	16s_sampl	P149_M3_	P149_M3_	subjectP14	M1	interventic	NR	R	R
Sample_29	16s_sampl	P150_M3_	P150_M3_	subjectP15	M1	interventic	R	NR	NR
Sample_29	16s_sampl	P16_M4_p	P16_M4_p	subject16	M1	interventic	NR	R	R
Sample_29	16s_sampl	P18_M4_p	P18_M4_p	subject18	M1	interventic	NR	NR	NR
Sample_29	16s_sampl	P19_M4_p	P19_M4_p	subject19	M1	interventic	NR	NR	R
Sample_30	16s_sampl	P1_M4_ph	P1_M4_ph	subject1	M1	interventic	NR	NR	NR
Sample_30	16s_sampl	P2_M4_ph	P2_M4_ph	subject2	M1	interventic	R	R	R
Sample_30	16s_sampl	P6_M4_ph	P6_M4_ph	subject6	M1	interventic	NR	R	R
Sample_30	16s_sampl	P7_M4_ph	P7_M4_ph	subject7	M1	interventic	R	R	R
Sample_30	16s_sampl	3R443_M5	3R443_M5	subject443	M2	interventic	NR	R	NR
Sample_30	16s_sampl	3R474_M5	3R474_M5	subject474	M2	interventic	NR	R	NR

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_3C	16s_sampl	3R478_M5	3R478_M5	subject478	M2	interventic	NR	NR	R
Sample_3C	16s_sampl	3R479_M5	3R479_M5	subject479	M2	interventic	NR	NR	R
Sample_3C	16s_sampl	3R481_M5	3R481_M5	subject481	M2	interventic	NR	NR	R
Sample_3C	16s_sampl	3R482_M5	3R482_M5	subject482	M2	interventic	NR	NR	NR
Sample_31	16s_sampl	3R486_M5	3R486_M5	subject486	M2	interventic	NR	R	R
Sample_31	16s_sampl	3R487_M5	3R487_M5	subject487	M2	interventic	R	R	R
Sample_31	16s_sampl	3R488_M5	3R488_M5	subject488	M2	interventic	R	R	NR
Sample_31	16s_sampl	3R492_M5	3R492_M5	subject492	M2	interventic	NR	R	NR
Sample_31	16s_sampl	3R493_M5	3R493_M5	subject493	M2	interventic	NR	R	R
Sample_31	16s_sampl	3R495_M5	3R495_M5	subject495	M2	interventic	NR	R	NR
Sample_31	16s_sampl	3R496_M5	3R496_M5	subject496	M2	interventic	NR	NR	NR
Sample_31	16s_sampl	3R506_M5	3R506_M5	subject506	M2	interventic	NR	NR	NR
Sample_31	16s_sampl	3R512_M5	3R512_M5	subject512	M2	interventic	NR	NR	R
Sample_31	16s_sampl	863_M2_L	863_M2_L	subject863	M2	interventic	NR	R	R
Sample_32	16s_sampl	P10_M5_p	P10_M5_p	subject10	M2	interventic	R	R	R
Sample_32	16s_sampl	P115_M2_	P115_M2_	subjectP11	M2	interventic	NR	R	R
Sample_32	16s_sampl	P11_M5_p	P11_M5_p	subject11	M2	interventic	NR	NR	R
Sample_32	16s_sampl	P121_M2_	P121_M2_	subjectP12	M2	interventic	NR	R	NR
Sample_32	16s_sampl	P125_M2_	P125_M2_	subjectP12	M2	interventic	NR	R	R
Sample_32	16s_sampl	P127_M2_	P127_M2_	subjectP12	M2	interventic	NR	R	R
Sample_32	16s_sampl	P129_M2_	P129_M2_	subjectP12	M2	interventic	NR	R	R
Sample_32	16s_sampl	P13_M4_p	P13_M4_p	subjectP13	M2	interventic	NR	NR	R
Sample_32	16s_sampl	P134_M2_	P134_M2_	subject13	M2	interventic	NR	R	R
Sample_32	16s_sampl	P140_M2_	P140_M2_	subjectP14	M2	interventic	NR	NR	NR
Sample_33	16s_sampl	P143_M2_	P143_M2_	subjectP14	M2	interventic	R	R	R
Sample_33	16s_sampl	P145_M2_	P145_M2_	subjectP14	M2	interventic	R	R	R
Sample_33	16s_sampl	P147_M2_	P147_M2_	subjectP14	M2	interventic	R	R	NR
Sample_33	16s_sampl	P149_M2_	P149_M2_	subjectP14	M2	interventic	NR	R	R
Sample_33	16s_sampl	P150_M2_	P150_M2_	subjectP15	M2	interventic	R	NR	NR
Sample_33	16s_sampl	P16_M5_p	P16_M5_p	subject16	M2	interventic	NR	R	R
Sample_33	16s_sampl	P18_M5_p	P18_M5_p	subject18	M2	interventic	NR	NR	NR
Sample_33	16s_sampl	P19_M5_p	P19_M5_p	subject19	M2	interventic	NR	NR	R
Sample_33	16s_sampl	P1_M5_ph	P1_M5_ph	subject1	M2	interventic	NR	NR	NR
Sample_33	16s_sampl	P2_M5_ph	P2_M5_ph	subject2	M2	interventic	R	R	R
Sample_34	16s_sampl	P6_M5_ph	P6_M5_ph	subject6	M2	interventic	NR	R	R
Sample_34	16s_sampl	P7_M5_ph	P7_M5_ph	subject7	M2	interventic	R	R	R
Sample_34	16s_sampl	3R443_M6	3R443_M6	subject443	M3	interventic	NR	R	NR
Sample_34	16s_sampl	3R474_M6	3R474_M6	subject474	M3	interventic	NR	R	NR
Sample_34	16s_sampl	3R478_M6	3R478_M6	subject478	M3	interventic	NR	NR	R
Sample_34	16s_sampl	3R479_M6	3R479_M6	subject479	M3	interventic	NR	NR	R
Sample_34	16s_sampl	3R481_M6	3R481_M6	subject481	M3	interventic	NR	NR	R
Sample_34	16s_sampl	3R482_M6	3R482_M6	subject482	M3	interventic	NR	NR	NR
Sample_34	16s_sampl	3R486_M6	3R486_M6	subject486	M3	interventic	NR	R	R
Sample_34	16s_sampl	3R487_M6	3R487_M6	subject487	M3	interventic	R	R	R
Sample_35	16s_sampl	3R488_M6	3R488_M6	subject488	M3	interventic	R	R	NR
Sample_35	16s_sampl	3R492_M6	3R492_M6	subject492	M3	interventic	NR	R	NR
Sample_35	16s_sampl	3R493_M6	3R493_M6	subject493	M3	interventic	NR	R	R
Sample_35	16s_sampl	3R495_M6	3R495_M6	subject495	M3	interventic	NR	R	NR
Sample_35	16s_sampl	3R496_M6	3R496_M6	subject496	M3	interventic	NR	NR	NR
Sample_35	16s_sampl	3R506_M6	3R506_M6	subject506	M3	interventic	NR	NR	NR
Sample_35	16s_sampl	3R512_M6	3R512_M6	subject512	M3	interventic	NR	NR	R

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_35	16s_sampl	863_M4_L	863_M4_L	subject863	M3	interventic	NR	R	R
Sample_35	16s_sampl	P10_M6_p	P10_M6_p	subject10	M3	interventic	R	R	R
Sample_35	16s_sampl	P115_M4_	P115_M4_	subjectP11	M3	interventic	NR	R	R
Sample_36	16s_sampl	P121_M4_	P121_M4_	subjectP12	M3	interventic	NR	R	NR
Sample_36	16s_sampl	P125_M4_	P125_M4_	subjectP12	M3	interventic	NR	R	R
Sample_36	16s_sampl	P127_M4_	P127_M4_	subjectP12	M3	interventic	NR	R	R
Sample_36	16s_sampl	P129_M4_	P129_M4_	subjectP12	M3	interventic	NR	R	R
Sample_36	16s_sampl	P13_M6_p	P13_M6_p	subjectP13	M3	interventic	NR	NR	R
Sample_36	16s_sampl	P134_M3_	P134_M3_	subject13	M3	interventic	NR	R	R
Sample_36	16s_sampl	P140_M4_	P140_M4_	subjectP14	M3	interventic	NR	NR	NR
Sample_36	16s_sampl	P143_M4_	P143_M4_	subjectP14	M3	interventic	R	R	R
Sample_36	16s_sampl	P145_M4_	P145_M4_	subjectP14	M3	interventic	R	R	R
Sample_36	16s_sampl	P147_M4_	P147_M4_	subjectP14	M3	interventic	R	R	NR
Sample_37	16s_sampl	P149_M4_	P149_M4_	subjectP14	M3	interventic	NR	R	R
Sample_37	16s_sampl	P150_M4_	P150_M4_	subjectP15	M3	interventic	R	NR	NR
Sample_37	16s_sampl	P16_M6_p	P16_M6_p	subject16	M3	interventic	NR	R	R
Sample_37	16s_sampl	P18_M6_p	P18_M6_p	subject18	M3	interventic	NR	NR	NR
Sample_37	16s_sampl	P19_M6_p	P19_M6_p	subject19	M3	interventic	NR	NR	R
Sample_37	16s_sampl	P1_M6_ph	P1_M6_ph	subject1	M3	interventic	NR	NR	NR
Sample_37	16s_sampl	P2_M6_ph	P2_M6_ph	subject2	M3	interventic	R	R	R
Sample_37	16s_sampl	P6_M6_ph	P6_M6_ph	subject6	M3	interventic	NR	R	R
Sample_37	16s_sampl	P7_M6_ph	P7_M6_ph	subject7	M3	interventic	R	R	R
Sample_37	16s_sampl	3R443_M7	3R443_M7	subject443	M4	interventic	NR	R	NR
Sample_38	16s_sampl	3R474_M7	3R474_M7	subject474	M4	interventic	NR	R	NR
Sample_38	16s_sampl	3R478_M7	3R478_M7	subject478	M4	interventic	NR	NR	R
Sample_38	16s_sampl	3R479_M7	3R479_M7	subject479	M4	interventic	NR	NR	R
Sample_38	16s_sampl	3R481_M7	3R481_M7	subject481	M4	interventic	NR	NR	R
Sample_38	16s_sampl	3R482_M7	3R482_M7	subject482	M4	interventic	NR	NR	NR
Sample_38	16s_sampl	3R486_M7	3R486_M7	subject486	M4	interventic	NR	R	R
Sample_38	16s_sampl	3R487_M7	3R487_M7	subject487	M4	interventic	R	R	R
Sample_38	16s_sampl	3R492_M7	3R492_M7	subject492	M4	interventic	NR	R	NR
Sample_38	16s_sampl	3R493_M7	3R493_M7	subject493	M4	interventic	NR	R	R
Sample_38	16s_sampl	3R495_M7	3R495_M7	subject495	M4	interventic	NR	R	NR
Sample_39	16s_sampl	3R496_M7	3R496_M7	subject496	M4	interventic	NR	NR	NR
Sample_39	16s_sampl	3R506_M7	3R506_M7	subject506	M4	interventic	NR	NR	NR
Sample_39	16s_sampl	3R512_M7	3R512_M7	subject512	M4	interventic	NR	NR	R
Sample_39	16s_sampl	P10_M7_p	P10_M7_p	subject10	M4	interventic	R	R	R
Sample_39	16s_sampl	P134_M4_	P134_M4_	subject13	M4	interventic	NR	R	R
Sample_39	16s_sampl	P16_M7_p	P16_M7_p	subject16	M4	interventic	NR	R	R
Sample_39	16s_sampl	P18_M7_p	P18_M7_p	subject18	M4	interventic	NR	NR	NR
Sample_39	16s_sampl	P1_M7_ph	P1_M7_ph	subject1	M4	interventic	NR	NR	NR
Sample_39	16s_sampl	P2_M7_ph	P2_M7_ph	subject2	M4	interventic	R	R	R
Sample_39	16s_sampl	P6_M7_ph	P6_M7_ph	subject6	M4	interventic	NR	R	R
Sample_40	16s_sampl	P7_M7_ph	P7_M7_ph	subject7	M4	interventic	R	R	R
Sample_40	16s_sampl	3R474_M8	3R474_M8	subject474	M5	interventic	NR	R	NR
Sample_40	16s_sampl	3R478_M8	3R478_M8	subject478	M5	interventic	NR	NR	R
Sample_40	16s_sampl	3R479_M8	3R479_M8	subject479	M5	interventic	NR	NR	R
Sample_40	16s_sampl	3R481_M8	3R481_M8	subject481	M5	interventic	NR	NR	R
Sample_40	16s_sampl	3R482_M8	3R482_M8	subject482	M5	interventic	NR	NR	NR
Sample_40	16s_sampl	3R487_M8	3R487_M8	subject487	M5	interventic	R	R	R
Sample_40	16s_sampl	3R488_M8	3R488_M8	subject488	M5	interventic	R	R	NR

medRxiv preprint doi: https://doi.org/10.1101/2022.10.28.22281668; this version posted January 18, 2023. 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.
All rights reserved. No reuse allowed without permission.

Sample_4C	16s_sampl	3R495_M8	3R495_M8	subject495	M5	interventic	NR	R	NR
Sample_4C	16s_sampl	3R496_M8	3R496_M8	subject496	M5	interventic	NR	NR	NR
Sample_41	16s_sampl	3R506_M8	3R506_M8	subject506	M5	interventic	NR	NR	NR
Sample_41	16s_sampl	3R512_M8	3R512_M8	subject512	M5	interventic	NR	NR	R
Sample_41	16s_sampl	P10_M8_p	P10_M8_p	subject10	M5	interventic	R	R	R
Sample_41	16s_sampl	P16_M8_p	P16_M8_p	subject16	M5	interventic	NR	R	R
Sample_41	16s_sampl	P18_M8_p	P18_M8_p	subject18	M5	interventic	NR	NR	NR
Sample_41	16s_sampl	P19_M8_p	P19_M8_p	subject19	M5	interventic	NR	NR	R
Sample_41	16s_sampl	P2_M8_ph	P2_M8_ph	subject2	M5	interventic	R	R	R
Sample_41	16s_sampl	P6_M8_ph	P6_M8_ph	subject6	M5	interventic	NR	R	R
Sample_41	16s_sampl	P7_M8_ph	P7_M8_ph	subject7	M5	interventic	R	R	R