Distinct smell and taste disorder phenotype of post-acute COVID-19 sequelae

Supplementary Material

HACT and CovILD study teams

Supplementary Methods

Study procedures and variables

The complete list of study variables and stratification scheme is provided in **Supplementary Table S1** for the HACT study and in **Supplementary Table S2** for the CovILD study.

COVID-19 symptoms

A total of 42 was surveyed in the HACT study cohorts (**Supplementary Figure S1** and **Supplementary Table S1**). The symptom duration was coded as follows: absent: 0 days, 1 - 3 days: 3 days, up to 1 week: 7 days, up to 2 weeks: 14 days, up to 4 weeks: 28 days, up to 3 months: 90 days, up to 6 months: 90 days, over 6 months: 90 days. Acute symptoms were defined as complaints present during the first 14 days after clinical onset of COVID-19. Long COVID symptoms were defined as complaints present for \geq 28 days. PASC symptoms were defined as manifestations present for \geq 90 days (1,2).

In the CovILD study, a total of 8 symptoms (reduced physical performance, hyposmia/anosmia, dyspnea, sleep problems, cough, fever, night sweating, gastrointestinal symptoms) were prospectively recorded with a standardized questionnaire at each of 60-, 100-, 180- and 360-day post COVID-19 follow-up (**Supplementary Figure S3** and **Supplementary Table S2**). Acute COVID-19 symptoms were assessed retrospectively (3,4).

Rating of physical recovery, mental health and quality of life in the HACT study

Self-perceived complete recovery, rehabilitation need and new medication since COVID-19 at the time of study participation were surveyed as single yes/no items. Percentage of physical performance loss as compared with the time before COVID-19 was rated with a 0 - 100% scale (5,6). Quality of life impairment (QoL) and overall mental health impairment (OMH) were rated with a 4 item Likert scale each (possible answers: "excellent," "good," "fair," "poor," scored: 0, 1, 2, and 3) (5,6). Anxiety (ANX) and depression (DPR) were assessed with the Patient Health Questionnaire (PHQ-4) (5–7). Psychosocial stress was scored with a modified 7 item PHQ stress module as described before (5,6,8).

Rating of hyposmia with sniffing stick test

Objective hyposmia at the 100-day and 360-day follow-up in the CovILD study participants was investigated with the 16-item sniffing stick test as described (9). Clinically relevant hyposmia was defined as \leq 12 correct answers (9,10). In the analysis, participants with the complete answers concerning self-reported hyposmia and complete test results were included.

Statistical analysis

Data transformation, descriptive statistic

Data transformation and statistical analysis was done with R version 4.0.5 with *tidyverse* data science environment (11). Analysis results were visualized with *ggplot2* (12), *cowplot* (13) as well as in-house developed *ExDA* (https://github.com/PiotrTymoszuk/ExDA) and *figur* (https://github.com/PiotrTymoszuk/figur) packages.

Descriptive statistics including median with interquartile ranges and frequency of complete answers for numeric and categorical variables were calculated with base R functions and *ExDA* package.

Statistical hypothesis testing

Since multiple study variables were non-normally distributed as assessed by Shapiro-Wilk test and visual assessment of their distribution (quantile-quantile plots), statistical significance for differences in outcome numeric variables were assessed with Mann-Whitney U test by Wilcoxon r effect size statistic (two groups) or Kruskal-Wallis test with η^2 effect size statistic. Differences in frequency distribution for categorical outcome variables were assessed by χ^2 test with Cramer V effect size statistic. Interrater assessment of self-reported and sniffing test hyposmia was accomplished with Cohen's κ statistic (14). κ significance ($\kappa \neq 0$) was estimated with Wald Z test. P values were adjusted for multiple testing with Benjamini-Hochberg method (15) separately for each analysis task and cohort. Packages *rstatix* (16), *vcd* (17) and *Exda* (https://github.com/PiotrTymoszuk/ExDA) were used for statistical hypothesis testing.

Modeling of symptom recovery kinetic

To model recovery kinetics for binary symptom variables (0: absent, 1: present), secondorder mixed-effect logistic (categorical features) modeling was applied (packages: *lme4*, *lmerTest* and development package *kinet* [https://github.com/PiotrTymoszuk/kinet]) (18– 20). Each model followed the general formula:

 $Response \sim time + time^2 + (1 \lor individual)$

where $(1 \lor individual)$ indicates the random effect of the individual and time and $time^2$ indicate the first- and second-order time effect terms. The first-order term estimate was interpreted as a measure of the recovery speed and the second-order term estimate was used to assess the plateau/rebound effect. Significance of the accuracy gain of the full second-order model compared with the nested null model was determined by likelihood ratio test (LRT) versus the nested first-order and null models, respectively. Likelihood ratio λ statistic (full versus null model) was used as an effect size measure. Individuals from the HACT or CovILD studies with the complete longitudinal symptom record were included in the kinetic modeling tasks. Results of the kinetic modeling were adjusted for multiple comparisons with Benjamini-Hochberg method (15).

Symptom-symptom distances and multi-dimensional scaling

To assess co-occurrence or exclusivity of symptoms, simple matching distances between manifestations of each acute COVID-19, long COVID and PASC in the HACT study cohorts were calculated (package *scrime* and development package *clustTools* [https://github.com/PiotrTymoszuk/clustTools]) (21,22). Subsequently, the distance matrix was subjected to multi-dimensional scaling (MDS, k = 2 dimensions, package *stats*, function *cmdscale()*). Association of specific symptoms was assessed by visual analysis of MDS coordinate plots.

Apriori analysis of acute COVID-19, long COVID and PASC symptoms in the HACT study

Frequent combinations of symptoms of long COVID and PASC in the HACT study cohorts were identified with the apriori algorithm (package *arules*) (23,24) with the minimal support cutoff of 0.1, 2 - 10 item transaction length, confidence > 0.8 and lift > 2. The support statistic were used to estimate the symptom combination frequency. The confidence value was treated as an estimate of conditional probability of the symptom co-occurrence. The lift statistic was interpreted as a measure of the symptom dependence (lift = 1, symptoms are independent). The complete analysis results are presented in **Supplementary Tables S3** and **S4**.

Clustering analysis

COVID-19 recovery clusters of the training Autria (AT) HACT cohort participants in respect to symptom-specific recovery times (Figure 1A) were defined with the PAM (partitioning around medoids) algorithm and Euclidean distance statistic (packages cluster, philentropy and development package *clustTools* [https://github.com/PiotrTymoszuk/clustTools]) (25,26). The set of participants with the complete clustering variable set was included in the analysis. The symptom recovery times were not subjected to any type of preprocessing. The choice of the clustering procedure was motivated by the analysis of the clustering variance (ratio of the total between-cluster to total sum of squares) and clustering structure stability in 10-fold cross-validation (metric: rate of correct cluster assignment, cluster assignment predicted by k = 5 nearest neighbors label propagation algorithm, package *clustTools*) (27,28) for several clustering algorithms as presented in **Supplementary Figure S8A**. The optimal number of clusters was determined by the bend of the total within-cluster sum of squares curve (Supplementary Figure S8B, package *factoextra*) (29). Permutation importance of specific clustering variables was investigated by calculating difference in clustering variance (ratio of total between-cluster sum of squares to total sum of squares) between the initial clustering object and the clustering object with the given variable reshuffled at random (package *clustTools*). Assignment of the Italy HACT cohort participants to the recovery clusters was accomplished with k-nearest neighbors label propagation algorithm (k = 5) (28). The clustering efficacy in the training AT cohort and the test IT cohort measured by clustering variance statistic defined above was similar (AT: 0.59, IT: 0.57).

Data and source code availability

The raw data files will be made available upon request. The entire analysis pipeline was published at https://github.com/PiotrTymoszuk/hyposmia_analsis_pipeline.

Supplementary Tables

Supplementary Table S1: HACT study variables.

Variable name	Variable label	Unit	Description
ID	Patient ID		patient ID
cohort	Cohort		Study cohort
acute_covid	Acute COVID-19 symptoms		Acute COVID-19 symptoms (first two weeks)
sex	Sex		participant's sex
age	Age	years	Participant's age at survey completion
bmi_class_before	BMI before COVID-19		Body mass index class before COVID-19
cohabitants	Household size	persons	Household size in persons
household_size	Household size		Household size class: single, pair or bigger
education_class	Education		Tertiary vs. non-tertiary education
employment_before	Employment status		Employment status before COVID-19
employment_sector	Employment sector		Employment sector
obs_time	Observation time		Observation time: test to survey completion
smoking	Smoking history		Smoking history before COVID-19
comorb_present	Comorbidity		At least one comorbidity present
comorb_sum	Sum of co-morbidities		Sum of the co-morbidities queried in the survey
hypertension	Hypertension		Hypertension before acute COVID-19
heart_circulation	Cardiovascular disease		Cardiovascular disease before acute COVID-19
diabetes	Diabetes		Diabetes before acute COVID-19
lung	Pulmonary disease		Pulmonary disease before COVID-19
gastrointenstinal	Gastrointestinal disease		Gastrointestinal disease before COVID-19
malignancy	Malignancy		Malignancy before acute COVID-19
hay_fever	Hay fever/allergy		Hay fever or allergy before acute COVID-19

Variable name	Variable label	Unit	Description
autoimmunity	Autoimmunity		Autoimmune disease before acute COVID-19
frequent_flu_like	Freq. resp. infections		Frequent upper respiratory tract infections (more than two per year) before COVID-19
two_plus_infections_a ntibiotics	Freq. bact. Infections		Frequent bacterial infections requiring antibiotic treatment (more than two per year) before COVID-19
depression_burnout	Pre-CoV depression/anxiety		Depression, anxiety or burnout before acute COVID-19
insomnia	Pre-CoV sleep disorders		Sleep disorders before COVID-19
night_dyspnoe	Sleep apnea		Night apnea before COVID-19
bruxism	Bruxism		Bruxism before COVID-19
pins_needles_feet	Feet paresthesia		Feet/leg paresthesia before acute COVID-19
daily_medication	Daily medication		Daily medication class, number of drugs taken daily before COVID-19
cov_outbreak	SARS-CoV2 outbreak		Infection during the spring 2020, summer/fall 2020, winter/spring 2021
illness_feeling	Severe illness feeling		Subjective feeling of acute infection
sum_symptoms_acute	# acute symptoms		Sum of acute symptoms (first two weeks)
sum_symptoms_long	# persistent symptoms		Sum of persistent symptoms (4 weeks and longer)
weight_loss_kg	Weight loss	kg	Weight loss during/after acute COVID-19
hair_loss	Hair loss		Hair loss during/after acute COVID-19
incomplete_covelesce nce	Incomplete recovery		Self-perceived incomplete convalescence
perf_impairment	Physical performance loss	percent	Percent loss of physical performance after COVID-19 as compared with the time before
new_medication_fup	New medication after COVID-19		New medication after COVID-19
rehabilitation_fup_ne eded	Subjective need for rehabilitation		Self-reported need for rehabilitation after COVID-19
phq_anxiety_score	ANX score		Anxiety score

stress_score Stress mental_health_score OM	R score ess score H score L score	 Depression score Stress score Overall mental health score
 mental_health_score OM	H score L score	
	L score	Overall mental health score
life_quality_score QoI		
		Quality of life score
fever Fev	ver	Fever
ague Shiv	vering	Shivering
sore_throat Sor	e throat	Sore throat
running_nose Rur	nning nose	Running nose
fatigue Fati	igue	Fatigue
dry_cough Dry	/ cough	Dry cough
wet_cough We	t cough	Wet cough
breath_short Tac	chypnea	Tachypnea
dyspnoe Dys	spnea	Dyspnea
chest_pain Che	est pain	Chest pain
tachycardia Tac	chycardia	Tachycardia
extrasystole Palj	pitations	Palpitations
joint_pain Join	nt pain	Joint pain
bone_pain Bor	ne pain	Bone pain
muscle_pain Mus	scle pain	Muscle pain
abdominal_pain Abc	dominal pain	Abd. pain
nausea Nau	usea	Nausea
vomiting Vor	niting	Vomiting
dim_appetite Dim	n. appetite	Dim. appetite
diarrhea Dia	rrhea	Diarrhea
dizziness Dizz	ziness	Dizziness
anosmia Hyr	po/anosmia	Hyposmia/anosmia

Variable name	Variable label	Unit	Description
taste_loss	Hypo/ageusia		Hypogeusia/ageusia
confusion	Confusion		Confusion
tingle_feet	Tingling feet		Tingling feet
tingle_hands	Tingling hands		Tingling hands
ache_feet	Burning feet		Burning feet
ache_hands	Burning hands		Burning hands
numb_feet	Numb feet		Numb feet
numb_hands	Numb hands		Numb hands
unhandiness_walk	Imp. walk		Imp. walk
unhandiness_microm otor	Imp. FMS		Imp. f. m. s.
sleep_prob	Sleeplessness		Sleeplessness
fatigue_day	Tiredness at day		Tiredness at day
imp_concentration	Imp. concentration		Imp. concentration
forgetfulness	Forgetfulness		Forgetfulness
swelling	Swelling		Swelling
blue_fingers	Blue fingers/toes		Blue fingers/toes
urticaria	Urticaria		Urticaria
blister_rash	Blistering rash		Blistering rash
net_rash	Marmorated skin		Bl. marm. skin.
red_eyes	Red eyes		Red eyes

Variable name	Variable label	Unit
ID		
time_numeric	Time post diagnosis	days
sex	Sex	
age	Age	years
weight_class	Weight class	
comorb_present	Comorbidity present	%
no_comorb	# comorbidities	
endometabolic_comorb	Metabolic disease	%
hypertension_comorb	Hypertension	%
cardiovascular_comorb	CVD	%
diabetes_comorb	Diabetes	%
pulmonary_comorb	Pulmonary disease	%
gastro_comorb	GID	%
malingancy_comorb	Malignancy	%
immdef_comorb	Immune deficiency	%
cat_WHO	COVID-19 severity	
sleep_sympt	Sleep problems	%
dyspnoe_sympt	Dyspnea	%
cough_sympt	Cough	%
fever_sympt	Fever	%
night_sweat_sympt	Night sweat	%
gastro_sympt	Gastrointestinal	%
anosmia_sympt	Hypo/anosmia	%
fatigue_sympt	Reduced performance	%

Supplementary Table S2: CovILD study variables.

Supplementary Table S3: Results of apriori analysis for long COVID-19 symptoms in the HACT study . Top 10 most frequent symptom combinations (highest support value) in the AT (Austria) and IT (Italy) cohort are presented. The full results are available as a supplementary Excel table.

Cohort	Symptoms	Support	Confidence	Lift
АТ	{Tachypnea, Tiredness at day,Imp. concentration} → {Fatigue}	0.184	0.946	2.10
АТ	{Fatigue, Hypo/anosmia} → {Hypo/ageusia}	0.192	0.836	2.34
AT	{Tachypnea, Imp. concentration} \rightarrow {Fatigue}	0.196	0.913	2.02
АТ	${Forgetfulness} → {Imp. concentration}$	0.200	0.857	2.67
АТ	{Hypo/ageusia, Tiredness at day} → {Hypo/anosmia}	0.215	0.920	2.29
AT	{Hypo/anosmia, Tiredness at day} → {Hypo/ageusia}	0.215	0.844	2.36
AT	{Fatigue, Imp. concentration} → {Tiredness at day}	0.251	0.945	2.05
AT	{Tachypnea, Tiredness at day} → {Fatigue}	0.251	0.902	2.00
AT	{Hypo/ageusia} → {Hypo/anosmia}	0.328	0.918	2.29
AT	{Hypo/anosmia} → {Hypo/ageusia}	0.328	0.818	2.29
IT	{Fatigue, Hypo/ageusia,Tiredness at day} → {Hypo/anosmia}	0.213	0.948	2.42
IT	{Fatigue, Hypo/anosmia,Tiredness at day} → {Hypo/ageusia}	0.213	0.929	2.57
IT	{Fatigue, Forgetfulness} → {Imp. concentration}	0.225	0.906	2.72
IT	{Tiredness at day, Forgetfulness} \rightarrow {Imp. concentration}	0.234	0.943	2.84
IT	{Fatigue, Hypo/ageusia} → {Hypo/anosmia}	0.237	0.944	2.41
IT	{Fatigue, Hypo/anosmia} → {Hypo/ageusia}	0.237	0.918	2.55
IT	{Hypo/ageusia, Tiredness at day} → {Hypo/anosmia}	0.239	0.944	2.41
IT	{Hypo/anosmia, Tiredness at day} → {Hypo/ageusia}	0.239	0.927	2.57
IT	{Forgetfulness} → {Imp. concentration}	0.253	0.878	2.64
IT	{Hypo/ageusia} → {Hypo/anosmia}	0.337	0.935	2.39
IT	{Hypo/anosmia} → {Hypo/ageusia}	0.337	0.862	2.39

Supplementary Table S4: Results of apriori analysis for PASC symptoms in the HACT study. Top 10 most frequent symptom combinations (highest support value) in the AT (Austria) and IT (Italy) cohort are presented. The full results are available as a supplementary Excel table.

Cohort	Symptoms	Support	Confidence	Lift
AT	{Fatigue, Forgetfulness} → {Tiredness at day}	0.113	0.947	2.95
AT	{Tiredness at day, Forgetfulness} \rightarrow {Fatigue}	0.113	0.844	3.16
AT	{Tiredness at day, Forgetfulness} \rightarrow {Imp. concentration}	0.121	0.906	4.06
AT	${Fatigue, Tachypnea} → {Tiredness at day}$	0.129	0.816	2.54
AT	{Tachypnea, Tiredness at day} → {Fatigue}	0.129	0.827	3.09
AT	${Forgetfulness} \rightarrow {Imp. concentration}$	0.144	0.802	3.59
AT	{Fatigue, Imp. concentration} → {Tiredness at day}	0.152	0.948	2.95
AT	{Tiredness at day, Imp. concentration} \rightarrow {Fatigue}	0.152	0.811	3.04
AT	{Imp. concentration} \rightarrow {Tiredness at day}	0.188	0.841	2.62
AT	{Hypo/ageusia} → {Hypo/anosmia}	0.198	0.880	2.93
IT	{Imp. concentration, Forgetfulness} \rightarrow {Tiredness at day}	0.185	0.823	2.51
IT	{Tiredness at day, Forgetfulness} \rightarrow {Imp. concentration}	0.185	0.963	3.61
IT	{Tiredness at day, Imp. concentration} \rightarrow {Forgetfulness}	0.185	0.859	3.27
IT	{Tiredness at day, Imp. concentration} \rightarrow {Fatigue}	0.199	0.924	2.76
IT	{Fatigue, Imp. concentration} → {Tiredness at day}	0.199	0.944	2.88
IT	{Hypo/ageusia} → {Hypo/anosmia}	0.201	0.956	3.55
IT	{Imp. concentration} \rightarrow {Tiredness at day}	0.215	0.807	2.46
IT	${Forgetfulness} \rightarrow {Imp. concentration}$	0.225	0.857	3.21
IT	{Imp. concentration} \rightarrow {Forgetfulness}	0.225	0.842	3.21
IT	{Tiredness at day} → {Fatigue}	0.274	0.836	2.50
IT	${Fatigue} \rightarrow {Tiredness at day}$	0.274	0.818	2.50

Supplementary Table S5: Demographic and baseline clinical characteristic at the COVID-19 onset of the HACT study participants assigned to the recovery clusters, Austria (AT) cohort.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
Sex	female: 79% (n = 77) male: 21% (n = 21) complete: n = 98	female: 57% (n = 140) male: 43% (n = 106) complete: n = 246	female: 76% (n = 103) male: 24% (n = 32) complete: n = 135	p < 0.001	V = 0.22
Age, years	median: 42 [IQR: 30 - 50] range: 21 - 80 complete: n = 98	median: 43 [IQR: 29 - 53] range: 18 - 77 complete: n = 246	median: 48 [IQR: 38 - 53] range: 21 - 70 complete: n = 135	p = 0.045	$\eta^2 = 0.012$
BMI before COVID- 19 ^b	normal: 62% (n = 60) overweight: 24% (n = 23) obesity: 14% (n = 14) complete: n = 97	normal: 55% (n = 133) overweight: 29% (n = 70) obesity: 17% (n = 41) complete: n = 244	normal: 47% (n = 64) overweight: 31% (n = 42) obesity: 21% (n = 29) complete: n = 135	ns (p = 0.39)	V = 0.073
Education	non-tertiary: 64% (n = 62) tertiary: 36% (n = 35) complete: n = 97	non-tertiary: 63% (n = 154) tertiary: 37% (n = 92) complete: n = 246	non-tertiary: 64% (n = 86) tertiary: 36% (n = 49) complete: n = 135	ns (p = 0.99)	V = 0.012
Employment status	employed: 87% (n = 85) unemployed: 7.1% (n = 7) leave: 3.1% (n = 3) retired: 3.1% (n = 3) complete: n = 98	employed: 80% (n = 198) unemployed: 9.3% (n = 23) leave: 1.6% (n = 4) retired: 8.5% (n = 21) complete: n = 246	employed: 85% (n = 115) unemployed: 7.4% (n = 10) leave: 0.74% (n = 1) retired: 6.7% (n = 9) complete: n = 135	ns (p = 0.56)	V = 0.079
Observation time	median: 180 [IQR: 130 - 210] range: 93 - 400 complete: n = 98	median: 190 [IQR: 130 - 220] range: 90 - 400 complete: n = 246	median: 180 [IQR: 140 - 220] range: 90 - 380 complete: n = 135	ns (p = 0.85)	$\eta^2 = -$ 0.0029
Comorbidity	46% (n = 45) complete: n = 98	44% (n = 109) complete: n = 246	61% (n = 83) complete: n = 135	p = 0.0095	V = 0.15
Hypertension	9.2% (n = 9) complete: n = 98	10% (n = 25) complete: n = 246	13% (n = 17) complete: n = 135	ns (p = 0.82)	V = 0.041
Cardiovascular disease	0% (n = 0) complete: n = 98	2.8% (n = 7) complete: n = 246	2.2% (n = 3) complete: n = 135	ns (p = 0.36)	V = 0.076

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
Diabetes	2% (n = 2) complete: n = 98	1.6% (n = 4) complete: n = 246	0.74% (n = 1) complete: n = 135	ns (p = 0.82)	V = 0.04
Pulmonary disease	0% (n = 0) complete: n = 98	4.5% (n = 11) complete: n = 246	5.2% (n = 7) complete: n = 135	ns (p = 0.14)	V = 0.1
Gastrointestinal disease	1% (n = 1) complete: n = 98	2% (n = 5) complete: n = 246	1.5% (n = 2) complete: n = 135	ns (p = 0.88)	V = 0.032
Malignancy	0% (n = 0) complete: n = 98	0.81% (n = 2) complete: n = 246	5.9% (n = 8) complete: n = 135	p = 0.0025	V = 0.17
Hay fever/allergy	13% (n = 13) complete: n = 98	17% (n = 41) complete: n = 246	25% (n = 34) complete: n = 135	ns (p = 0.073)	V = 0.12
Autoimmunity ^c	7.1% (n = 7) complete: n = 98	4.1% (n = 10) complete: n = 246	11% (n = 15) complete: n = 135	ns (p = 0.056)	V = 0.12
Freq. resp. infections ^d	5.1% (n = 5) complete: n = 98	5.3% (n = 13) complete: n = 246	10% (n = 14) complete: n = 135	ns (p = 0.2)	V = 0.093
Freq. bact. Infections	1% (n = 1) complete: n = 98	3.7% (n = 9) complete: n = 246	9.6% (n = 13) complete: n = 135	p = 0.01	V = 0.15
Pre-CoV depression/anxiety	6.1% (n = 6) complete: n = 98	2.8% (n = 7) complete: n = 246	9.6% (n = 13) complete: n = 135	p = 0.038	V = 0.13
Pre-CoV sleep disorders	5.1% (n = 5) complete: n = 98	2.4% (n = 6) complete: n = 246	4.4% (n = 6) complete: n = 135	ns (p = 0.53)	V = 0.063
Daily medication	absent: 66% (n = 65) 1 - 4 drugs: 30% (n = 29) 5 drugs and more: 4.1% (n = 4) complete: n = 98	absent: 66% (n = 162) 1 - 4 drugs: 33% (n = 80) 5 drugs and more: 1.6% (n = 4) complete: n = 246	absent: 50% (n = 68) 1 - 4 drugs: 49% (n = 66) 5 drugs and more: 0.74% (n = 1) complete: n = 135	p = 0.0092	V = 0.13

^aCategorical variables: χ^2 test with Cramer V effect size statistic. Numeric variables: Kruskal-Wallis test with η^2 effect size statistic. P values corrected form multiple testing with Benjamini-Hochberg method.

^bBMI: body mass index, overweight > 25 kg/m², obesity > 30 kg/m²,

^cFrequent respiratory infections, > 2 per year.

^dFrequent bacterial infections with antibiotic therapy, > 2 per year.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
Sex	female: 83% (n = 54) male: 17% (n = 11) complete: n = 65	female: 63% (n = 150) male: 37% (n = 87) complete: n = 237	female: 77% (n = 96) male: 23% (n = 29) complete: n = 125	p = 0.0032	V = 0.18
Age, years	median: 46 [IQR: 34 - 55] range: 18 - 71 complete: n = 65	median: 43 [IQR: 32 - 53] range: 18 - 77 complete: n = 237	median: 48 [IQR: 38 - 56] range: 19 - 95 complete: n = 125	p = 0.024	$\eta^2 = 0.016$
BMI before COVID- 19 ^b	normal: 82% (n = 53) overweight: 12% (n = 8) obesity: 6.2% (n = 4) complete: n = 65	normal: 67% (n = 155) overweight: 26% (n = 61) obesity: 6.5% (n = 15) complete: n = 231	normal: 57% (n = 70) overweight: 28% (n = 35) obesity: 15% (n = 18) complete: n = 123	p = 0.0077	V = 0.14
Education	non-tertiary: 57% (n = 37) tertiary: 43% (n = 28) complete: n = 65	non-tertiary: 57% (n = 135) tertiary: 43% (n = 102) complete: n = 237	non-tertiary: 62% (n = 78) tertiary: 38% (n = 47) complete: n = 125	ns (p = 0.6)	V = 0.05
Employment status	employed: 80% (n = 52) unemployed: 11% (n = 7) leave: 3.1% (n = 2) retired: 6.2% (n = 4) complete: n = 65	employed: 79% (n = 188) unemployed: 11% (n = 27) leave: 2.5% (n = 6) retired: 6.8% (n = 16) complete: n = 237	employed: 86% (n = 108) unemployed: 4.8% (n = 6) leave: 0% (n = 0) retired: 8.8% (n = 11) complete: n = 125	ns (p = 0.27)	V = 0.1
Observation time	median: 140 [IQR: 120 - 280] range: 92 - 370 complete: n = 65	median: 140 [IQR: 110 - 260] range: 90 - 390 complete: n = 237	median: 140 [IQR: 120 - 300] range: 90 - 380 complete: n = 125	ns (p = 0.21)	$\eta^2 = 0.0045$
Comorbidity	35% (n = 23) complete: n = 65	38% (n = 89) complete: n = 237	58% (n = 73) complete: n = 125	p < 0.001	V = 0.2
Hypertension	7.7% (n = 5) complete: n = 65	6.8% (n = 16) complete: n = 237	12% (n = 15) complete: n = 125	ns (p = 0.28)	V = 0.083
Cardiovascular disease	0% (n = 0) complete: n = 65	3.4% (n = 8) complete: n = 237	4% (n = 5) complete: n = 125	ns (p = 0.34)	V = 0.077

Supplementary Table S6: Demographic and baseline clinical characteristic at the COVID-19 onset of the HACT study participants assigned to the recovery clusters, Italy (IT) cohort.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
Diabetes	0% (n = 0) complete: n = 65	0% (n = 0) complete: n = 237	0.8% (n = 1) complete: n = 125	ns (p = 0.34)	V = 0.075
Pulmonary disease	3.1% (n = 2) complete: n = 65	2.1% (n = 5) complete: n = 237	4% (n = 5) complete: n = 125	ns (p = 0.6)	V = 0.051
Gastrointestinal disease	0% (n = 0) complete: n = 65	0.42% (n = 1) complete: n = 237	1.6% (n = 2) complete: n = 125	ns (p = 0.37)	V = 0.071
Malignancy	6.2% (n = 4) complete: n = 65	2.5% (n = 6) complete: n = 237	5.6% (n = 7) complete: n = 125	ns (p = 0.28)	V = 0.083
Hay fever/allergy	7.7% (n = 5) complete: n = 65	11% (n = 27) complete: n = 237	15% (n = 19) complete: n = 125	ns (p = 0.34)	V = 0.076
Autoimmunity ^c	6.2% (n = 4) complete: n = 65	4.6% (n = 11) complete: n = 237	9.6% (n = 12) complete: n = 125	ns (p = 0.25)	V = 0.089
Freq. resp. infections ^d	0% (n = 0) complete: n = 65	1.3% (n = 3) complete: n = 237	8.8% (n = 11) complete: n = 125	p < 0.001	V = 0.2
Freq. bact. Infections	0% (n = 0) complete: n = 65	0.42% (n = 1) complete: n = 237	3.2% (n = 4) complete: n = 125	ns (p = 0.071)	V = 0.12
Pre-CoV depression/anxiety	4.6% (n = 3) complete: n = 65	3% (n = 7) complete: n = 237	9.6% (n = 12) complete: n = 125	p = 0.044	V = 0.13
Pre-CoV sleep disorders	3.1% (n = 2) complete: n = 65	1.7% (n = 4) complete: n = 237	11% (n = 14) complete: n = 125	p < 0.001	V = 0.2
Daily medication	absent: 75% (n = 49) 1 - 4 drugs: 25% (n = 16) 5 drugs and more: 0% (n = 0) complete: n = 65	absent: 81% (n = 191) 1 - 4 drugs: 19% (n = 45) 5 drugs and more: 0.42% (n = 1) complete: n = 237	absent: 62% (n = 77) 1 - 4 drugs: 36% (n = 45) 5 drugs and more: 2.4% (n = 3) complete: n = 125	p = 0.0032	V = 0.14

^aCategorical variables: χ^2 test with Cramer V effect size statistic. Numeric variables: Kruskal-Wallis test with η^2 effect size statistic. P values corrected form multiple testing with Benjamini-Hochberg method.

^bBMI: body mass index, overweight > 25 kg/m², obesity > 30 kg/m²,

^cFrequent respiratory infections, > 2 per year.

^dFrequent bacterial infections with antibiotic therapy, > 2 per year.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
SARS-CoV2 outbreak	spring 2020: 55% (n = 54) summer/fall 2020: 43% (n = 42) winter/spring 2021: 2% (n = 2) complete: n = 98	spring 2020: 63% (n = 156) summer/fall 2020: 35% (n = 87) winter/spring 2021: 1.2% (n = 3) complete: n = 246	spring 2020: 53% (n = 71) summer/fall 2020: 47% (n = 64) winter/spring 2021: 0% (n = 0) complete: n = 135	ns (p = 0.16)	V = 0.09
Severe illness feeling	1 week: 52% (n = 51) 2 weeks: 26% (n = 25) 3 weeks: 8.2% (n = 8) 4 weeks: 8.2% (n = 8) over 4 weeks: 6.1% (n = 6) complete: n = 98	1 week: 69% (n = 170) 2 weeks: 19% (n = 46) 3 weeks: 6.9% (n = 17) 4 weeks: 4.1% (n = 10) over 4 weeks: 1.2% (n = 3) complete: n = 246	1 week: 24% (n = 33) 2 weeks: 22% (n = 30) 3 weeks: 14% (n = 19) 4 weeks: 12% (n = 16) over 4 weeks: 27% (n = 37) complete: n = 135	p < 0.001	V = 0.34
# acute symptoms ^b	median: 15 [IQR: 11 - 18] range: 2 - 30 complete: n = 98	median: 11 [IQR: 8 - 15] range: 1 - 28 complete: n = 246	median: 20 [IQR: 16 - 23] range: 4 - 42 complete: n = 135	p < 0.001	$\eta^2 = 0.26$
# persistent symptoms ^c	median: 3 [IQR: 2 - 5] range: 1 - 12 complete: n = 98	median: 0 [IQR: 0 - 1] range: 0 - 6 complete: n = 246	median: 7 [IQR: 4 - 12] range: 1 - 34 complete: n = 135	p < 0.001	$\eta^{2} = 0.72$
Weight loss, kg	median: 0.5 [IQR: 0 - 3] range: 0 - 8 complete: n = 98	median: 0 [IQR: 0 - 2.1] range: 0 - 11 complete: n = 244	median: 2 [IQR: 0 - 4.5] range: 0 - 15 complete: n = 134	p < 0.001	$\eta^2 = 0.039$
Hair loss	19% (n = 19) complete: n = 98	9.3% (n = 23) complete: n = 246	30% (n = 41) complete: n = 135	p < 0.001	V = 0.24
Incomplete recovery ^d	62% (n = 61) complete: n = 98	22% (n = 55) complete: n = 245	73% (n = 98) complete: n = 135	p < 0.001	V = 0.47
Physical performance loss, percent	median: 10 [IQR: 4 - 25] range: 0 - 69 complete: n = 98	median: 3.5 [IQR: 0 - 14] range: 0 - 100 complete: n = 244	median: 25 [IQR: 15 - 42] range: 0 - 92 complete: n = 134	p < 0.001	$\eta^2 = 0.26$

Supplementary Table S7: COVID-19 course and recovery in the HACT study participants assigned to the recovery clusters, Austria (AT) cohort.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a	
New medication after COVID-19	7.1% (n = 7) complete: n = 98	7.3% (n = 18) complete: n = 246	24% (n = 32) complete: n = 135	p < 0.001	V = 0.23	
Subjective need for rehabilitation	13% (n = 13) complete: n = 98	6.5% (n = 16) complete: n = 246	42% (n = 56) complete: n = 134	p < 0.001	V = 0.4	
ANX score ^e	median: 0 [IQR: 0 - 2] range: 0 - 5 complete: n = 98	median: 0 [IQR: 0 - 1] range: 0 - 6 complete: n = 246	median: 1.5 [IQR: 0 - 2] range: 0 - 6 complete: n = 134	p < 0.001	$\eta^2 = 0.11$	
DPR score ^f	median: 1 [IQR: 0 - 2] range: 0 - 6 complete: n = 98	median: 0 [IQR: 0 - 2] range: 0 - 6 complete: n = 246	median: 2 [IQR: 1 - 3] range: 0 - 6 complete: n = 135	p < 0.001	$\eta^{2} = 0.15$	
Stress score	median: 3.5 [IQR: 2 - 6] range: 0 - 19 complete: n = 98	median: 3 [IQR: 1 - 5] range: 0 - 16 complete: n = 246	median: 5 [IQR: 3 - 9] range: 0 - 16 complete: n = 135	p < 0.001	$\eta^2 = 0.064$	
OMH score ^g	median: 1 [IQR: 0 - 1] range: 0 - 3 complete: n = 98	median: 1 [IQR: 0 - 1] range: 0 - 3 complete: n = 246	median: 1 [IQR: 1 - 2] range: 0 - 3 complete: n = 135	p < 0.001	$\eta^2 = 0.072$	
QoL score ^h	median: 1 [IQR: 1 - 1] range: 0 - 3 complete: n = 98	median: 1 [IQR: 0 - 1] range: 0 - 3 complete: n = 246	median: 1 [IQR: 1 - 2] range: 0 - 3 complete: n = 135	p < 0.001	$\eta^2 = 0.052$	

 a Categorical variables: χ^{2} test with Cramer V effect size statistic. Numeric variables: Kruskal-Wallis test with η^{2} effect size statistic. P values corrected form multiple testing with Benjamini-Hochberg method.

^bNumber of acute COVID-19 symptoms.

^cNumber of long COVID symptoms.

^dSelf-reported incomplete recovery.

^eANX: anxiety.

^fDPR: depression.

^gOMH score: overall mental health impairment score

^hQoL score: quality of life impairment score

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a	
SARS-CoV2 outbreak	spring 2020: 34% (n = 22) summer/fall 2020: 65% (n = 42) winter/spring 2021: 1.5% (n = 1) complete: n = 65	$ \begin{array}{l} = 67) & = 40) \\ \text{er/fall 2020:} & \text{summer/fall 2020:} & \text{summer/fall 2020:} \\ \text{n} = 42) & 71\% \text{ (n} = 169) & 68\% \text{ (n} = 85) \\ \text{/spring 2021:} & \text{winter/spring 2021:} & \text{winter/spring 2021:} \\ \text{n} = 1) & 0.42\% \text{ (n} = 1) & 0\% \text{ (n} = 0) \end{array} $		ns (p = 0.55)	V = 0.062	
Severe illness feeling	1 week: 43% (n = 28) 2 weeks: 25% (n = 16) 3 weeks: 12% (n = 8) 4 weeks: 11% (n = 7) over 4 weeks: 9.2% (n = 6) complete: n = 65	1 week: 67% (n = 157) 2 weeks: 15% (n = 36) 3 weeks: 8.1% (n = 19) 4 weeks: 6.8% (n = 16) over 4 weeks: 3% (n = 7) complete: n = 235	1 week: 20% (n = 25) 2 weeks: 18% (n = 23) 3 weeks: 16% (n = 20) 4 weeks: 13% (n = 16) over 4 weeks: 33% (n = 41) complete: n = 125	p < 0.001	V = 0.35	
# acute symptoms ^b	median: 14 [IQR: 9 - 19] range: 4 - 31 complete: n = 65	9] 15] 26] ange: 4 - 31 range: 2 - 30 range: 8 - 38		p < 0.001	$\eta^{2} = 0.3$	
# persistent symptoms ^c	median: 2 [IQR: 2 - 4] range: 1 - 11 complete: n = 65	ange: 1 - 11 range: 0 - 7 range: 2 - 27		p < 0.001	$\eta^2 = 0.67$	
Weight loss, kg	median: 0 [IQR: 0 - 2] range: 0 - 5 complete: n = 65	median: 0 [IQR: 0 - 2] range: 0 - 8 complete: n = 236	median: 2 [IQR: 0 - 4] range: 0 - 15 complete: n = 124	p < 0.001	$\eta^2 = 0.049$	
Hair loss	25% (n = 16) complete: n = 65	9.7% (n = 23) complete: n = 237	31% (n = 39) complete: n = 125 p < 0.001		V = 0.25	
Incomplete recovery ^d	55% (n = 35) complete: n = 64	18% (n = 43) complete: n = 235	65% (n = 81) complete: n = 125	p < 0.001	V = 0.45	
Physical performance loss, percent	ce loss, median: 10 [IQR: 1 - median: 20] 16] range: 0 - 58 range: 0 complete: n = 65 complete		median: 29 [IQR: 20 - 50] range: 0 - 93 complete: n = 124	p < 0.001	$\eta^2 = 0.29$	

Supplementary Table S8: COVID-19 course and recovery in the HACT study participants assigned to the recovery clusters, Italy (AT) cohort.

Variable	Cluster #1	Cluster #2	Cluster #3	Significance ^a	Effect size ^a
New medication after COVID-19	13% (n = 8) complete: n = 63	8.2% (n = 19) complete: n = 231	20% (n = 25) complete: n = 123	p = 0.0092	V = 0.16
Subjective need for rehabilitation	17% (n = 11) complete: n = 65	3.8% (n = 9) complete: n = 235	34% (n = 42) complete: n = 124	p < 0.001	V = 0.37
ANX score ^e	median: 1 [IQR: 0 - 2] range: 0 - 6 complete: n = 65	median: 0 [IQR: 0 - 2] range: 0 - 6 complete: n = 237	median: 2 [IQR: 1 - 4] range: 0 - 6 complete: n = 125	p < 0.001	$\eta^2 = 0.14$
DPR score ^f	median: 2 [IQR: 0 - 2] range: 0 - 6 complete: n = 65	median: 1 [IQR: 0 - 2] range: 0 - 6 complete: n = 236	median: 2 [IQR: 2 - 4] range: 0 - 6 complete: n = 125	p < 0.001	$\eta^2 = 0.16$
Stress score	median: 4 [IQR: 1 - 6] range: 0 - 13 complete: n = 65	median: 3 [IQR: 1 - 6] range: 0 - 14 complete: n = 235	median: 6 [IQR: 4 - 9] range: 0 - 15 complete: n = 125	p < 0.001	$\eta^2 = 0.11$
OMH score ^g	median: 1 [IQR: 0 - 1] range: 0 - 2 complete: n = 65	median: 1 [IQR: 0 - 1] range: 0 - 3 complete: n = 237	median: 1 [IQR: 1 - 2] range: 0 - 3 complete: n = 125	p < 0.001	$\eta^{2} = 0.1$
QoL score ^h	median: 1 [IQR: 1 - 1] range: 0 - 3 complete: n = 65	median: 1 [IQR: 0 - 1] range: 0 - 3 complete: n = 237	median: 1 [IQR: 1 - 2] range: 0 - 3 complete: n = 125	p < 0.001	$\eta^2 = 0.095$

^aCategorical variables: χ^2 test with Cramer V effect size statistic. Numeric variables: Kruskal-Wallis test with η^2 effect size statistic. P values corrected form multiple testing with Benjamini-Hochberg method.

^bNumber of acute COVID-19 symptoms.

^cNumber of long COVID symptoms.

^dSelf-reported incomplete recovery.

^eANX: anxiety.

^fDPR: depression.

^gOMH score: overall mental health impairment score

^hQoL score: quality of life impairment score

Supplementary Figures

Α

Symptom frequency AT, HACT study, % cohort

В

Symptom frequency

IT, HACT study, % cohort

Fever-	O 56%	• 3.8%	• 0.63%	Fever-	071%	• 3%	• 0.7%
Sore throat -	Ó 50%	• 4%	• 1.7%	Shivering -	Õ 51%	• 2.3%	• 1.2%
Running nose -	O 43%	• 2.7%	• 1.3%	Sore throat -	O 43%	• 4.2%	• 1.2%
Shivering -	O 40%	• 0.63%	• 0.42%	Running nose -	O 40%	• 5.2%	• 1.6%
	Q 42 /0	0.0370	0.4270		Ψ 4 1/0	÷ 5.270	1.070
Tachypnea-	0 59%	O 37%	• 25%	Tachypnea-	0 53%	• 30%	• 21%
Dry cough -	0 56%	• 17%	• 5.8%	Dry cough	○ 49%	• 17%	• 5.4%
Dyspnea -	○ 38%	• 17%	• 11%	Dyspnea -	O 31%	• 14%	• 7.5%
Wet cough -	• 18%	• 4%	• 1.7%	Wet cough -	• 13%	• 3%	• 0.47
Joint pain	66%	• 15%	• 9.4%	Joint pain -	73%	0 22%	• 12%
Chest pain	0 49%	• 21%	• 14%	Muscle pain	68%	• 22%	• 12%
Muscle pain	T		• 9%	Bone pain			
	0 58%	• 15%			0 64%	• 18%	• 9.6%
Bone pain -	○ 44%	• 12%	• 7.5%	Chest pain	O 42%	• 17%	• 7.7%
Abdominal pain -	• 22%	• 5.8%	• 2.3%	Abdominal pain -	○ 26%	• 7.3%	• 2.8%
Dim. appetite -	64%	• 11%	• 2.3%	Dim. appetite -	061%	• 11%	• 2.8%
Diarrhea -	Ō 31%	• 4.8%	• 2.7%	Diarrhea -	O 34%	• 4.9%	• 1.6%
Nausea -	0 27%	• 5.6%	• 2.1%	Nausea -	O 31%	• 4.7%	• 2.3%
Vomiting -	• 6.1%	• 0.63%	• 0.21%	Vomiting-	• 8.7%	• 0.7%	• 0.23
Tachycardia	0.000/	0.400/	- 110/	Tachycardia	0.00%	- 100/	
	O 33%	• 16%	• 11%		• 26%	• 13%	• 8.2%
Palpitations -	• 17%	• 11%	• 8.6%	Palpitations -	• 17%	• 9.4%	• 7.3%
Red eyes -	• 19%	• 6.9%	• 3.5%	Red eyes	O 31%	• 12%	• 6.1%
Swelling -	• 4.4%	• 2.7%	• 2.5%	Swelling -	• 6.3%	• 4.7%	• 4%
Urticaria -	• 4.8%	• 1.7%	• 1.3%	Urticaria -	• 8%	• 3.7%	• 3%
Blistering rash -	• 3.3%	• 1.7%	• 1%	Blistering rash -	• 5.4%	• 3%	• 1.6%
Marmorated skin -	• 2.1%	• 1.5%	• 1.3%	Blue fingers/toes -	• 2.1%	• 1.6%	• 1.2%
Blue fingers/toes -	• 1.3%	• 0.84%	• 0.84%	Marmorated skin -	• 1.9%	• 1.4%	• 0.94
Hypo/anosmia -	70%	O 40%	• 30%	Hypo/anosmia ∔	75%	O 39%	0 27%
					¥		
Hypo/ageusia -	68%	O 36%	• 23%	Hypo/ageusia -	74%	O 36%	O 21%
Fatigue -	94%	○ 45%	• 27%	Fatigue -	92%	0 54%	0 33%
Tiredness at day -	84%	0 46%	0 32%	Tiredness at day -	7 9%	O 49%	0 33%
Sleeplessness -	Ŏ 39%	• 19%	• 14%	Sleeplessness -	Ŏ 31%	• 16%	• 11%
np. concentration	O 50%	O 32%	• 22%	Imp. concentration	Ó 49%	O 33%	0 27%
Forgetfulness	0 32%	0 23%	• 18%	Forgetfulness	0 35%	0 29%	0 26%
Dizziness -				Confusion -			• 11%
Confusion -	○ 46% ● 18%	• 11% • 9.2%	• 6.9% • 5.6%	Dizziness -	● 26% ● 31%	• 17% • 12%	• 5.6%
Tingling feet	• 13%	• 5.2%	• 4.8%	Imp. walk -	• 15%	• 9.8%	• 6.6%
Imp. walk	• 12%	• 4.6%	• 2.5%	Tingling hands	• 13%	• 8.4%	• 6.8%
Tingling hands	• 8.6%	• 3.5%	• 2.9%	Tingling feet	• 13%	• 6.8%	• 4.9%
Numb feet -	• 6.1%	• 3.3%	• 2.9%	Numb hands -	• 9.4%	• 7%	• 5.6%
Burning feet -	• 5.8%	• 2.9%	• 2.9%	Numb feet -	• 10%	• 6.6%	• 5.2%
Numb hands -	• 4.8%	• 3.1%	• 2.9%	Burning feet -	• 7%	• 4.2%	• 3.5%
Imp. FMS -	• 4.6%	• 1%	• 0.63%	Burning hands -	• 4.7%	• 3%	• 2.19
Burning hands	• 1.7%	• 1.5%	• 1.5%	Imp. FMS	• 4.4%	• 2.8%	• 2.3%
				· L			

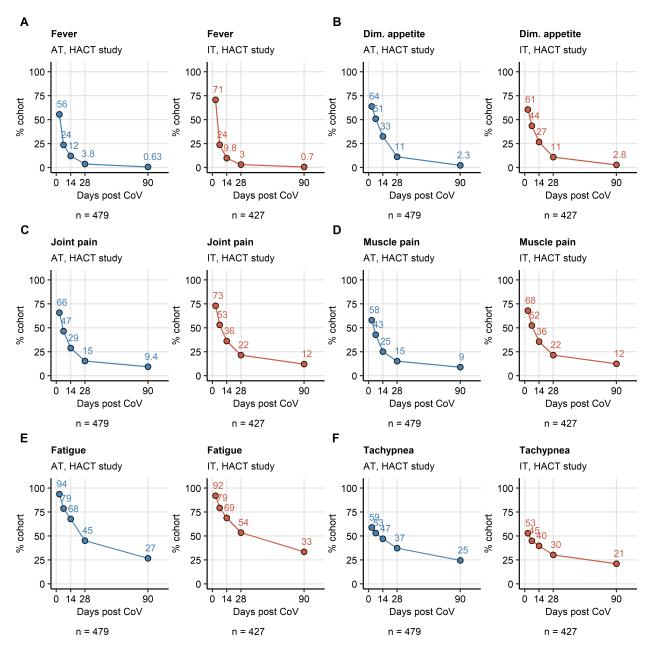
n = 479

n = 427

Supplementary Figure S1. Frequency of COVID-19 symptoms in acute COVID-19, long COVID and PASC in the HACT study.

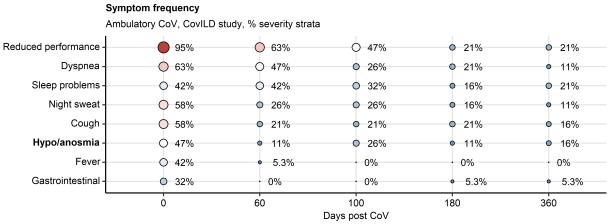
Frequency of symptoms of acute COVID-19 (first 14 days post clinical onset), long COVID (\geq 28 days) and PASC (post-acute sequelae of COVID-19, \geq 90 days) in the AT (Austria, **A**) and IT (Italy, **B**) HACT study cohorts expressed as percentages of the cohort. Point size and color represents the percentage. Numbers of complete observations are indicated under the plot.

Dim. appetite: diminished appetite, Imp. concentration: impaired concentration, Imp. walk: impaired walk, Imp. FMS: impaired fine motor skills.

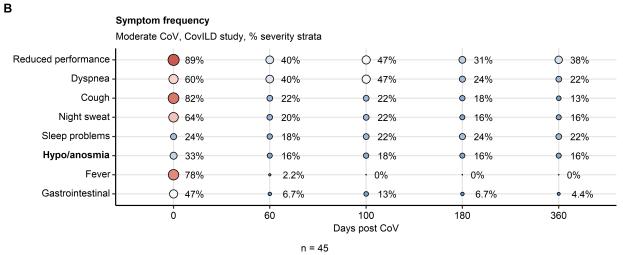


Supplementary Figure S2. Kinetic of recovery from leading acute COVID-19 symptoms in the HACT study.

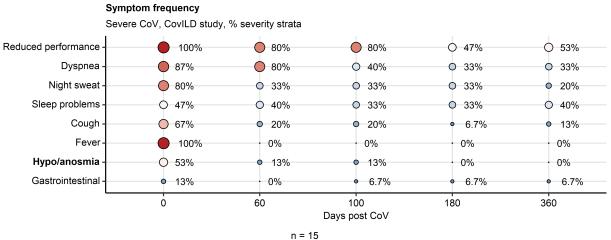
Percentages of individuals with fever (**A**), diminished appetite (**B**), joint pain (**C**), muscle pain (**D**), fatigue (**E**) and tachypnea (**F**) in the AT (Austria) and IT (Italy) HACT study cohorts at particular time points after clinical onset. Numbers of complete observations are indicated under the plots.









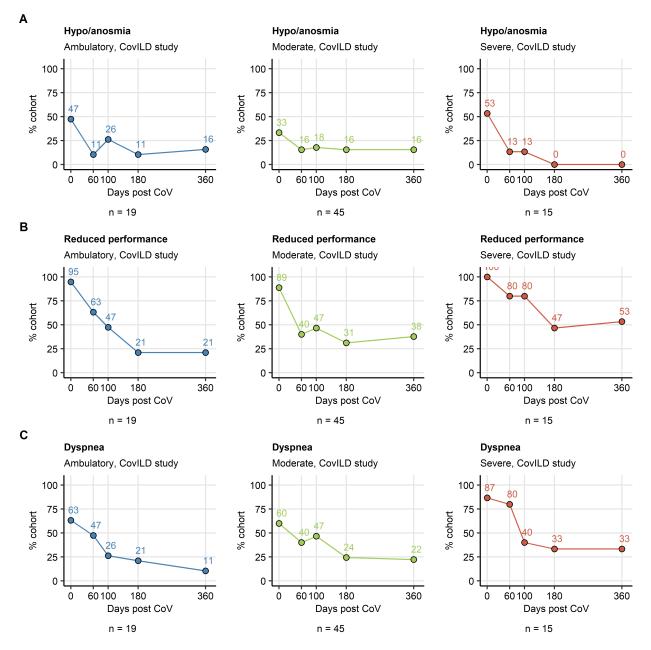


25

Α

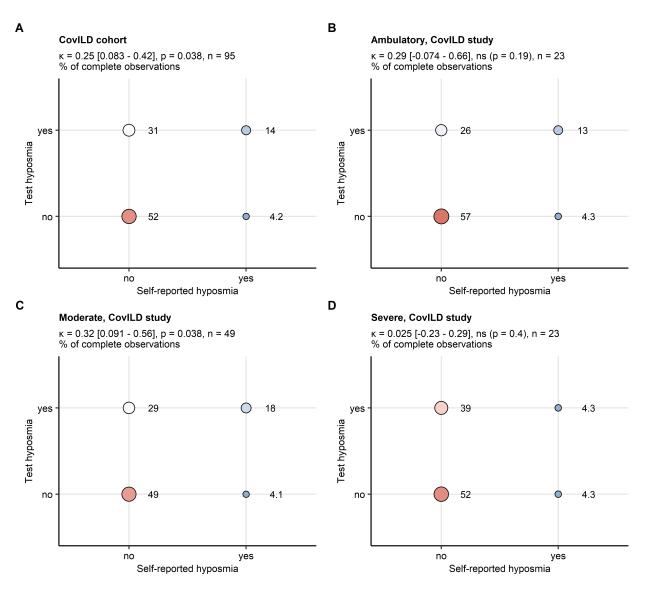
Supplementary Figure S3. Symptom frequency in ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

Frequency of symptoms during acute COVID-19 and at the 60-, 100-, 180- and 360-day follow-ups in ambulatory (**A**), moderate (**B**) and severe COVID-19 (**C**) participants of the CovILD study expressed as percentages of individuals with the complete longitudinal data set. Point size and color represents the percentage. Numbers of complete observations are indicated under the plots.



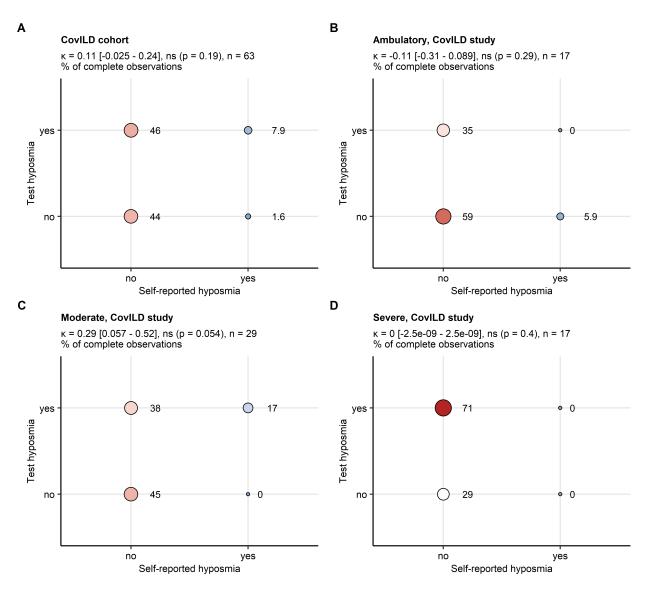
Supplementary Figure S4. Kinetic of recovery from smell disorders, reduced performance and dyspnea in ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

Percentages of individuals with the complete longitudinal data set suffering from smell disorders (**A**), reduced physical performance (**B**) and dyspnea (**C**) in the ambulatory, moderate and severe COVID-19 subsets of the CovILD study during acute COVID-19 and at the 60-, 100-, 180- and 360-day follow-ups. Numbers of complete observations are indicated under the plots.



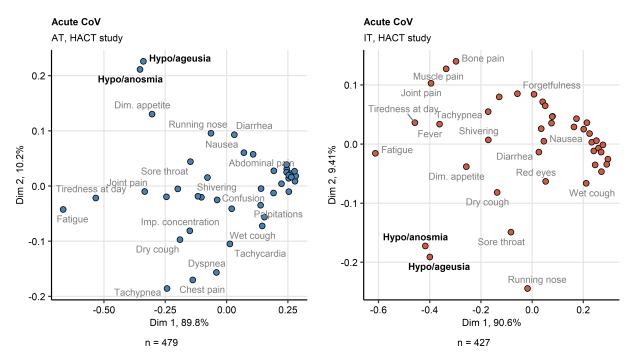
Supplementary Figure S5. Rates of self-reported hyposmia and hyposmia in sniffing stick test at 3-month post COVID-19 follow-up in the ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

Association of self-reported and sniffing stick test hyposmia rates was investigated with Cohen's κ statistic. Statistical significance was assessed with Wald's Z test corrected for multiple testing with Benjamini-Hochberg method. Percentages of individuals with self-reported and test hyposmia within the entire cohort (**A**), the ambulatory (**B**), moderate (**C**) and severe (**D**) COVID-19 subsets of the CovILD-19 subsets are presented in bubble plots. Point size and color represents the percentage. κ values with 95% confidence intervals, p values and numbers of complete observations are indicated in the plot captions.



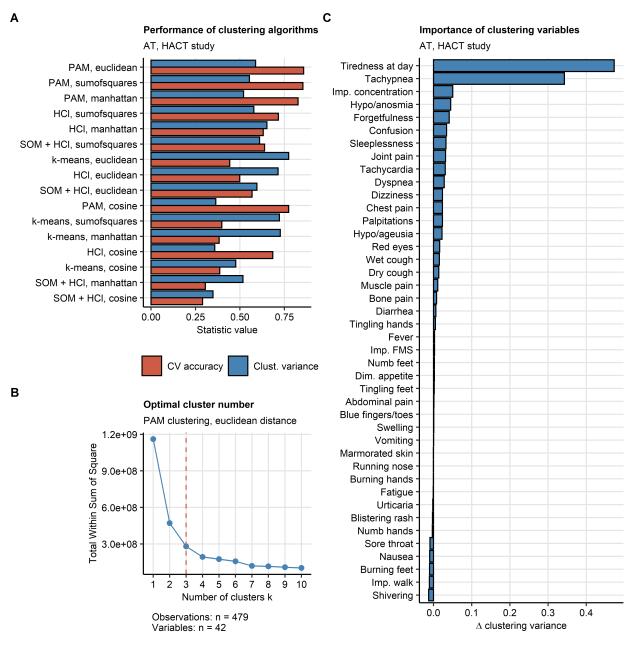
Supplementary Figure S6. Rates of self-reported hyposmia and hyposmia in sniffing stick test at 1-year post COVID-19 follow-up in the ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

Association of self-reported and sniffing stick test hyposmia rates was investigated with Cohen's κ statistic. Statistical significance was assessed with Wald's Z test corrected for multiple testing with Benjamini-Hochberg method. Percentages of individuals with self-reported and test hyposmia within the entire cohort (**A**), the ambulatory (**B**), moderate (**C**) and severe (**D**) COVID-19 subsets of the CovILD-19 subsets are presented in bubble plots. Point size and color represents the percentage. κ values with 95% confidence intervals, p values and numbers of complete observations are indicated in the plot captions.



Supplementary Figure S7. Multi-dimensional scaling analysis of acute COVID-19 symptoms in the HACT study.

Binary symptom occurrence data for acute COVID-19 (first 14 days after clinical onset) in the HACT study Austria (AT) and Italy (IT) cohorts were subjected to two-dimensional multi-dimensional scaling (MDS) with simple matching distance (SMD) between the symptoms. MDS coordinates are presented in point plots. Selected data points are labeled with the symptom names. Percentages of the data set variance associated with the MDS dimensions are indicated in the plot axes. Numbers of complete observations are indicated under the plots.



Supplementary Figure S8. Definition of the COVID-19 recovery clusters and clustering feature importance in the HACT study.

Individuals of the training Austria (AT) cohort of the HACT study were subjected to clustering in respect to symptom-specific recovery times with the PAM (partitioning around medoids) algorithm and Euclidean distance measure.

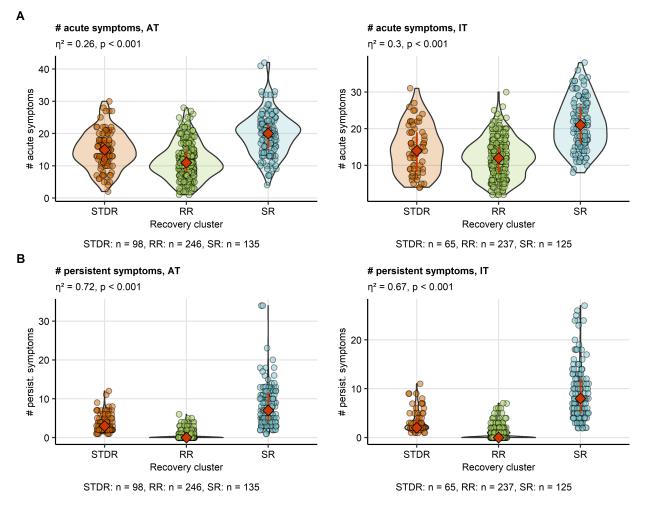
(A) Comparison of performance of various algorithms (HCl: hierarchical clustering, SOM + HCl: combined self-organizing map and hierarchical clustering, k-means) and distance statistic in clustering of the training data set investigated by clustering variance (ratio of

31

total between-cluster sum of squares to total sum of squares) and cluster assignment accuracy in 10-fold cross-validation (CV).

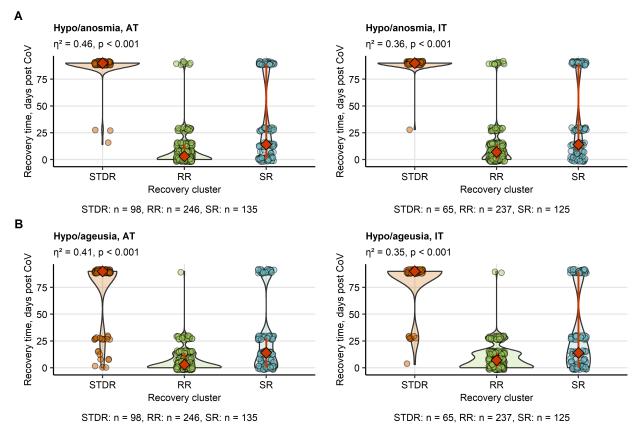
(B) Determination of the optimal cluster number in the PAM clustering of the training cohort by the bend of the total within-cluster sum of squares curve.

(C) Permutation importance of the clustering features (symptoms) for clustering of the training cohort expressed as the difference in clustering variance (ratio of total betweencluster sum of squares to total sum of squares) between the initial clustering object and the clustering object with the given variable reshuffled at random.



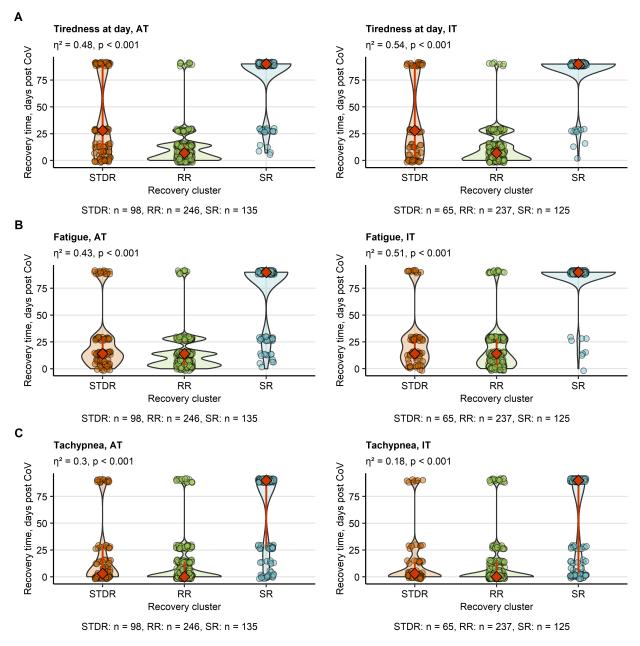
Supplementary Figure S9. Numbers of acute COVID-19 and long COVID symptoms in the HACT study recovery clusters.

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, Figure 5). Differences in numbers of acute COVID-19 (first 14 days after clinical onset, **A**) and long COVID symptoms (\geq 28 days, **B**) between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. Symptom counts are presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.



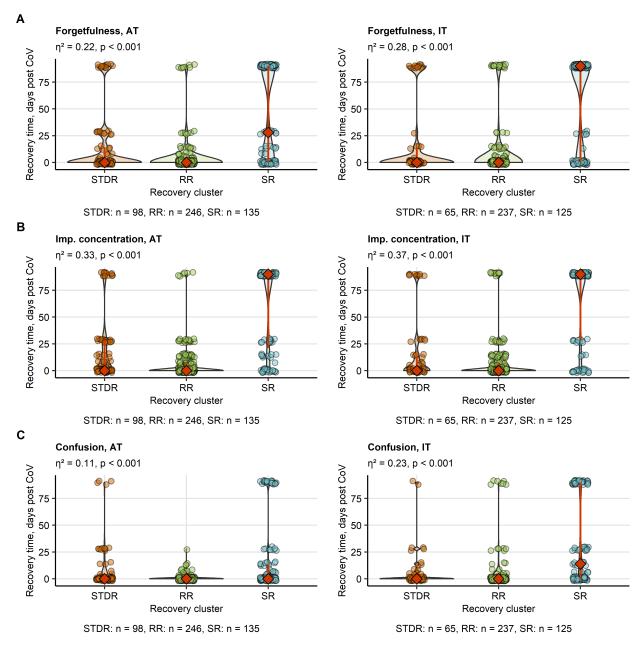
Supplementary Figure S10. Recovery from smell and taste disorders in the HACT study recovery clusters

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, Figure 5). Differences in recovery times for smell (**A**) and taste disorders (**D**) between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. Symptom counts are presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.



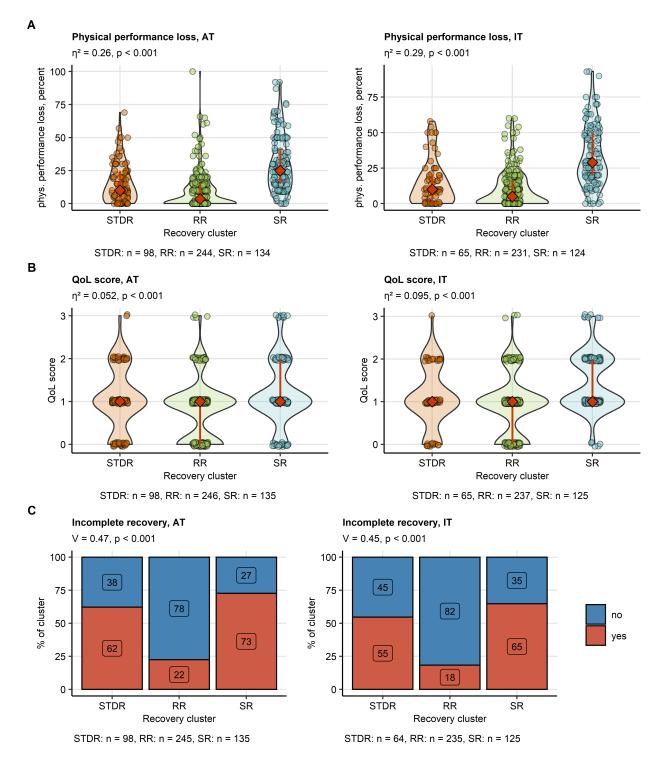
Supplementary Figure S11. Recovery from fatigue, tiredness and tachypnea in the HACT study recovery clusters.

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, Figure 5). Differences in recovery times for fatigue (**A**), day tiredness (**B**) and tachypnea (**C**) between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. Symptom counts are presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.



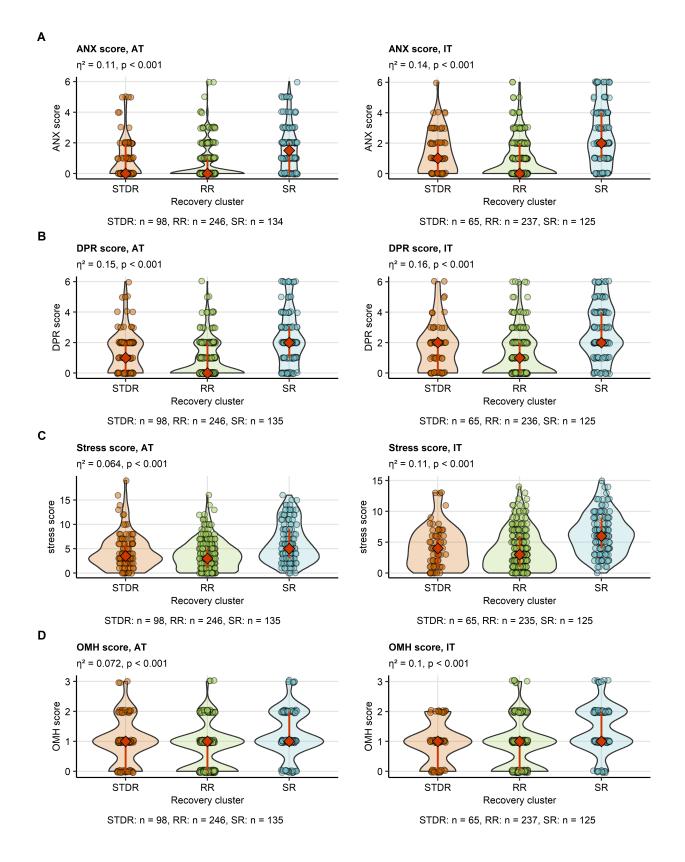
Supplementary Figure S12. Recovery from neurocognitive symptoms of COVID-19 in the HACT study recovery clusters.

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, **Figure 5**). Differences in recovery times for forgetfulness (**A**), concentration impairment (**B**) and confusion (**C**) between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. Symptom counts are presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.



Supplementary Figure S13. Physical performance loss, impairment of quality of life and self-perceived complete convalescence at the survey completion in the HACT study recovery clusters.

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, **Figure 5**). Differences in self-reported physical performance loss (**A**) and scoring of impairment of quality of life (QoL, **B**) at the survey completion between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. The scoring data were presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Differences in frequency of self-perceived complete convalescence (**C**) between the clusters were investigated by χ^2 test and Cramer V effect size statistic. The frequencies are presented as bar plots. P values were corrected for multiple testing with Benjamini-Hochberg method. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.



Supplementary Figure S14. Rating of anxiety, depression, overall mental health impairment and stress at the survey completion in the HACT study recovery clusters.

The recovery clusters were defined in the training Austria cohort (AT, **Supplementary Figure S8**) and the cluster assignment predicted in the test Italy cohort (IT, Figure 5). Differences in rating of anxiety (ANX, **A**), depression (DPR, **B**), stress (**C**) and self-reported overall mental health impairment (OMH, **D**) at the survey completion between the clusters were assessed by Kruskal-Wallis test and η^2 effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. The scoring data were presented in violin plots. Points represent single observations, orange diamonds with whiskers code for medians and interquartile ranges. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.

References

1. Venkatesan P. NICE guideline on long COVID. *The Lancet Respiratory medicine* (2021) 9:129. doi: 10.1016/S2213-2600(21)00031-X

2. Shah W, Hillman T, Playford ED, Hishmeh L. Managing the long term effects of covid-19: Summary of NICE, SIGN, and RCGP rapid guideline. *The BMJ* (2021) 372: doi: 10.1136/bmj.n136

3. Sonnweber T, Sahanic S, Pizzini A, Luger A, Schwabl C, Sonnweber B, Kurz K, Koppelstätter S, Haschka D, Petzer V, et al. Cardiopulmonary recovery after COVID-19 - an observational prospective multi-center trial. *The European respiratory journal* (2020) doi: 10.1183/13993003.03481-2020

4. Sonnweber T, Tymoszuk P, Sahanic S, Boehm A, Pizzini A, Luger A, Schwabl C, Nairz M, Grubwieser P, Kurz K, et al. Investigating phenotypes of pulmonary COVID-19 recovery - a longitudinal observational prospective multicenter trial. *eLife* (2022) 11: doi: 10.7554/ELIFE.72500

5. Hüfner K, Tymoszuk P, Ausserhofer D, Sahanic S, Pizzini A, Rass V, Galffy M, Böhm A, Kurz K, Sonnweber T, et al. Who Is at Risk of Poor Mental Health Following Coronavirus Disease-19 Outpatient Management? *Frontiers in medicine* (2022) 9: doi: 10.3389/FMED.2022.792881

6. Sahanic S, Tymoszuk P, Ausserhofer D, Rass V, Pizzini A, Nordmeyer G, Hüfner K, Kurz K, Weber PM, Sonnweber T, et al. Phenotyping of acute and persistent COVID-19 features in the outpatient setting: exploratory analysis of an international cross-sectional online survey. *Clinical Infectious Diseases* (2021) doi: 10.1093/CID/CIAB978

7. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, Schneider A, Brähler E. A 4item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders* (2010) 122:86–95. doi: 10.1016/j.jad.2009.06.019

8. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". Ergebnisse der Deutschen validierungsstudie. *Diagnostica* (2004) 50:171–181. doi: 10.1026/0012-1924.50.4.171

9. Rass V, Beer R, Schiefecker AJ, Kofler M, Lindner A, Mahlknecht P, Heim B, Limmert V, Sahanic S, Pizzini A, et al. Neurological outcome and quality of life 3 months after COVID-19: A prospective observational cohort study. *European Journal of Neurology* (2021) doi: 10.1111/ene.14803 10. Hummel T, Kobal G, Gudziol H, Mackay-Sim A. Normative data for the "Sniffin' Sticks" including tests of odor identification, odor discrimination, and olfactory thresholds: an upgrade based on a group of more than 3,000 subjects. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* (2007) 264:237–243. doi: 10.1007/S00405-006-0173-0

11. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the Tidyverse. *Journal of Open Source Software* (2019) 4:1686. doi: 10.21105/joss.01686

12. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. 1st ed. New York: Springer-Verlag (2016). pp. https://ggplot2.tidyverse.org

13. Wilke CO. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. 1st ed. Sebastopol: O'Reilly Media (2019). pp.

14. Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin* (1969) 72:323–327. doi: 10.1037/h0028106

15. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

16. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2021) https://cran.r-project.org/package=rstatix

17. Meyer D, Zeileis A, Hornik K. vcd: Visualizing Categorical Data. (2021) https://cran.r-project.org/web/packages/vcd/index.html

18. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* (2015) 67:1–48. doi: 10.18637/jss.v067.i01

19. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* (2017) 82:1–26. doi: 10.18637/jss.v082.i13

20. Box GE, Hunter SJ, Hunter WG. Statistics for experimenters: an introduction to design, data analysis, and model building. (2005)

21. Boriah S, Chandola V, Kumar V. Similarity measures for categorical data: A comparative evaluation. In. *Society for industrial and applied mathematics - 8th siam international conference on data mining 2008, proceedings in applied mathematics 130*. (2008). pp. 243–254 doi: 10.1137/1.9781611972788.22

22. Schwender H, Fritsch A. scrime: Analysis of High-Dimensional Categorical Data Such as SNP Data. (2018) https://cran.r-project.org/web/packages/scrime/index.html

23. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. *ACM SIGMOD Record* (1993) 22:207–216. doi: 10.1145/170036.170072

24. Hahsler M, Grün B, Hornik K. Arules - A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software* (2005) 14: doi: 10.18637/JSS.V014.I15

25. Drost H-G. Philentropy: Information Theory and Distance Quantification with R. *Journal of Open Source Software* (2018) 3:765. doi: 10.21105/joss.00765

26. Schubert E, Rousseeuw PJ. Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms. In. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*. Springer (2019). pp. 171–187 doi: 10.1007/978-3-030-32047-8_16

27. Lange T, Roth V, Braun ML, Buhmann JM. Stability-Based Validation of Clustering Solutions. *Neural Computation* (2004) 16:1299–1323. doi: 10.1162/089976604773717621

28. Leng M, Wang J, Cheng J, Zhou H, Chen X. Adaptive semi-supervised clustering algorithm with label propagation. *Journal of Software Engineering* (2014) 8:14–22. doi: 10.3923/JSE.2014.14.22

29. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. (2020) https://cran.r-project.org/web/packages/factoextra/index.html