

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

Supplementary Material

Hralth after COVID-19 in Tyrol and CovILD study teams

2022-07-12

Supplementary Methods

Study procedures and variables

The complete list of study variables and stratification scheme is provided in **Supplementary Table S1** for the Health after COVID-19 in Tyrol survey study and in **Supplementary Table S2** for the CovILD cohort.

COVID-19 symptoms

A total of 42 was surveyed in the survey 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: 30 months, up to 6 months: 3 months, over 6 months: 3 months. Acute symptoms were defined as complaints present during the first 14 days after clinical onset of COVID-19.

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 (1,2).

Rating of physical recovery, mental health and quality of life in the survey 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 (3,4). 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) (3,4). Anxiety (ANX) and depression (DPR) were assessed with the Patient Health Questionnaire (PHQ-4) (3-5). Psychosocial stress was scored with a modified 7 item PHQ stress module as described before (3,4,6).

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 (7). Clinically relevant

hyposmia was defined as ≤ 12 correct answers (7,8). 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 (9). Analysis results were visualized with *ggplot2* (10), *cowplot* (11) 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 (12). κ significance ($\kappa \neq 0$) was estimated with Wald Z test. P values were adjusted for multiple testing with Benjamini-Hochberg method (13) separately for each analysis task and cohort. Packages *rstatix* (14), *vcg* (15) 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), second-order mixed-effect logistic (categorical features) modeling was applied (packages: *lme4*, *lmerTest* and development package *kinet* [<https://github.com/PiotrTymoszuk/kinet>]) (16–18). Each model followed the general formula:

$$\text{Response} \sim \text{time} + \text{time}^2 + (1 \vee \text{individual})$$

where $(1 \vee \text{individual})$ 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 survey study or the CovILD cohort 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 (13).

Symptom-symptom distances and multi-dimensional scaling

To assess co-occurrence or exclusivity of symptoms, simple matching distances between manifestations during the first 14 days, at ≥ 28 days and at ≥ 3 months after clinical onset in the survey study cohorts were calculated (package *scrime* and development package *clustTools* [<https://github.com/PiotrTymoszuk/clustTools>]) (19,20). 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 COVID-19 symptoms in the survey study

Frequent combinations of symptoms during the first 14 days, at ≥ 28 days and at ≥ 3 months after clinical onset in the survey study cohorts were identified with the apriori algorithm (package *arules*) (21,22) 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).

Clustering analysis

COVID-19 recovery clusters of the training Autria (AT) survey 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>]) (23,24). 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 pre-processing. 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*) (25,26) 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*) (27). 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 survey cohort participants to the recovery clusters was accomplished with k-nearest neighbors label propagation algorithm ($k = 5$) (26). 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: Survey 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
gastrointestinal	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
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

Variable name	Variable label	Unit	Description
two_plus_infections_antibiotics	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_covelescence	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_needed	Subjective need for rehabilitation		Self-reported need for rehabilitation after COVID-19
phq_anxiety_score	ANX score		Anxiety score
phq_depression_score	DPR score		Depression score
stress_score	Stress score		Stress score
mental_health_score	OMH score		Overall mental health score

Variable name	Variable label	Unit	Description
life_quality_score	QoL score		Quality of life score
fever	Fever		Fever
ague	Shivering		Shivering
sore_throat	Sore throat		Sore throat
running_nose	Running nose		Running nose
fatigue	Fatigue		Fatigue
dry_cough	Dry cough		Dry cough
wet_cough	Wet cough		Wet cough
breath_short	Tachypnea		Tachypnea
dyspnoe	Dyspnea		Dyspnea
chest_pain	Chest pain		Chest pain
tachycardia	Tachycardia		Tachycardia
extrasystole	Palpitations		Palpitations
joint_pain	Joint pain		Joint pain
bone_pain	Bone pain		Bone pain
muscle_pain	Muscle pain		Muscle pain
abdominal_pain	Abdominal pain		Abd. pain
nausea	Nausea		Nausea
vomiting	Vomiting		Vomiting
dim_appetite	Dim. appetite		Dim. appetite
diarrhea	Diarrhea		Diarrhea
dizziness	Dizziness		Dizziness
anosmia	Hypo/anosmia		Hyposmia/anosmia
taste_loss	Hypo/ageusia		Hypogeusia/ageusia
confusion	Confusion		Confusion
tingle_feet	Tingling feet		Tingling feet

Variable name	Variable label	Unit	Description
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_micromotor	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

Supplementary Table S2: CovILD study variables.

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

Variable	STDR	RR	SR	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	η^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)	η^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
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

Variable	STDR	RR	SR	Significance ^a	Effect size ^a
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 from 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.

Supplementary Table S4: Demographic and baseline clinical characteristic at the COVID-19 onset of the survey study participants assigned to the recovery clusters, Italy (IT) cohort.

Variable	STDR	RR	SR	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
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

Variable	STDR	RR	SR	Significance ^a	Effect size ^a
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 from 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.

Supplementary Table S5: COVID-19 course and recovery in the survey study participants assigned to the recovery clusters, Austria (AT) cohort.

Variable	STDR	RR	SR	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
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	$\eta^2 = 0.26$
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	$\eta^2 = 0.072$

Variable	STDR	RR	SR	Significance ^a	Effect size ^a
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	$\eta^2 = 0.052$

^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 from multiple testing with Benjamini-Hochberg method.

^dSelf-reported incomplete recovery.

^eANX: anxiety.

^fDPR: depression.

^gOMH score: overall mental health impairment score

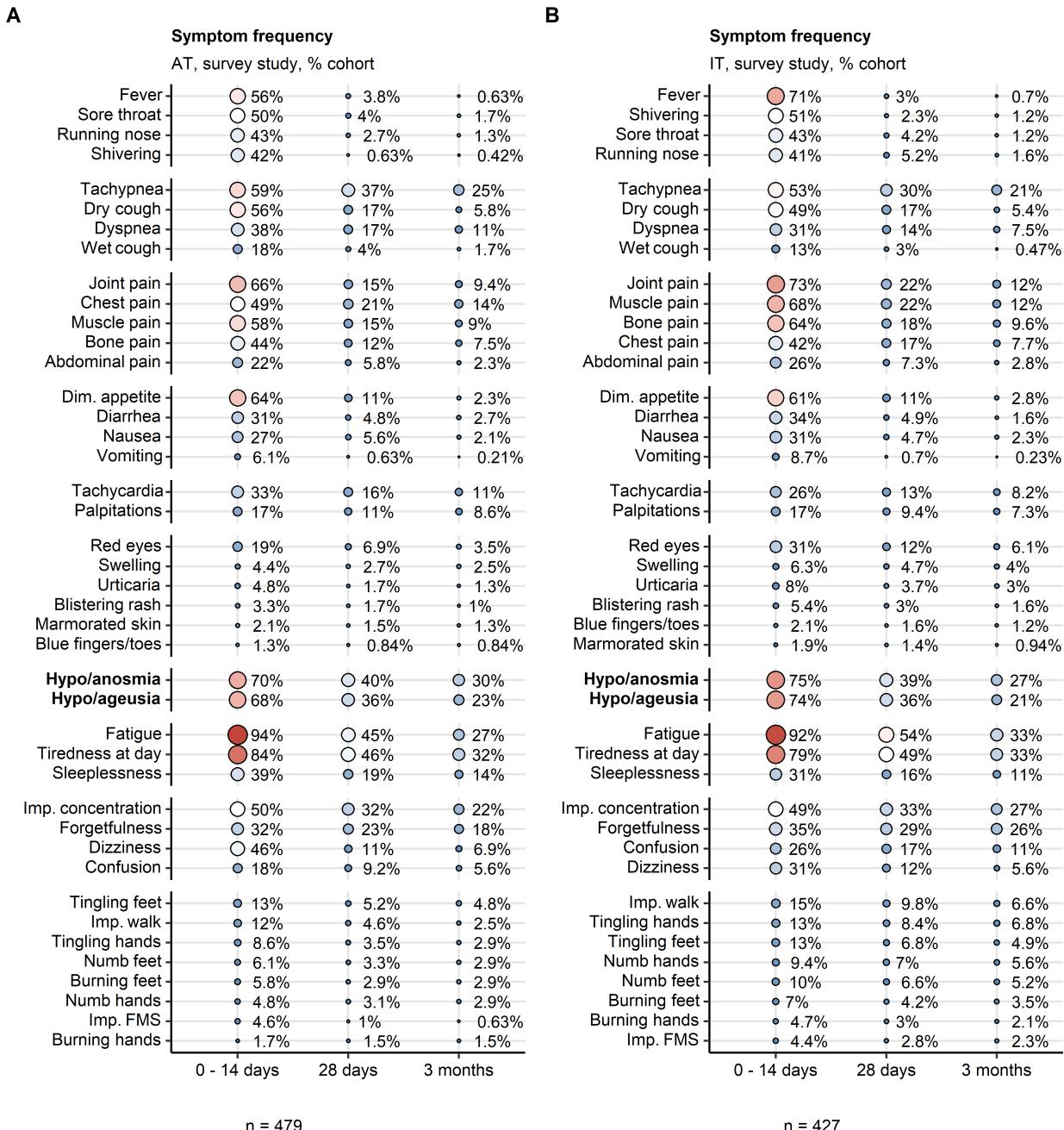
^hQoL score: quality of life impairment score

Supplementary Table S6: COVID-19 course and recovery in the survey study participants assigned to the recovery clusters, Italy (AT) cohort.

Variable	STDR	RR	SR	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	spring 2020: 28% (n = 67) summer/fall 2020: 71% (n = 169) winter/spring 2021: 0.42% (n = 1) complete: n = 237	spring 2020: 32% (n = 40) summer/fall 2020: 68% (n = 85) winter/spring 2021: 0% (n = 0) complete: n = 125	ns (p = 0.55)	V = 0.062
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	median: 10 [IQR: 1 - 20] range: 0 - 58 complete: n = 65	median: 5 [IQR: 0 - 16] range: 0 - 60 complete: n = 231	median: 29 [IQR: 20 - 50] range: 0 - 93 complete: n = 124	p < 0.001	$\eta^2 = 0.29$
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	$\eta^2 = 0.095$

Variable	STDR	RR	SR	Significance ^a	Effect size ^a
^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 from multiple testing with Benjamini-Hochberg method.					
^d Self-reported incomplete recovery.					
^e ANX: anxiety.					
^f DPR: depression.					
^g OMH score: overall mental health impairment score					
^h QoL score: quality of life impairment score					

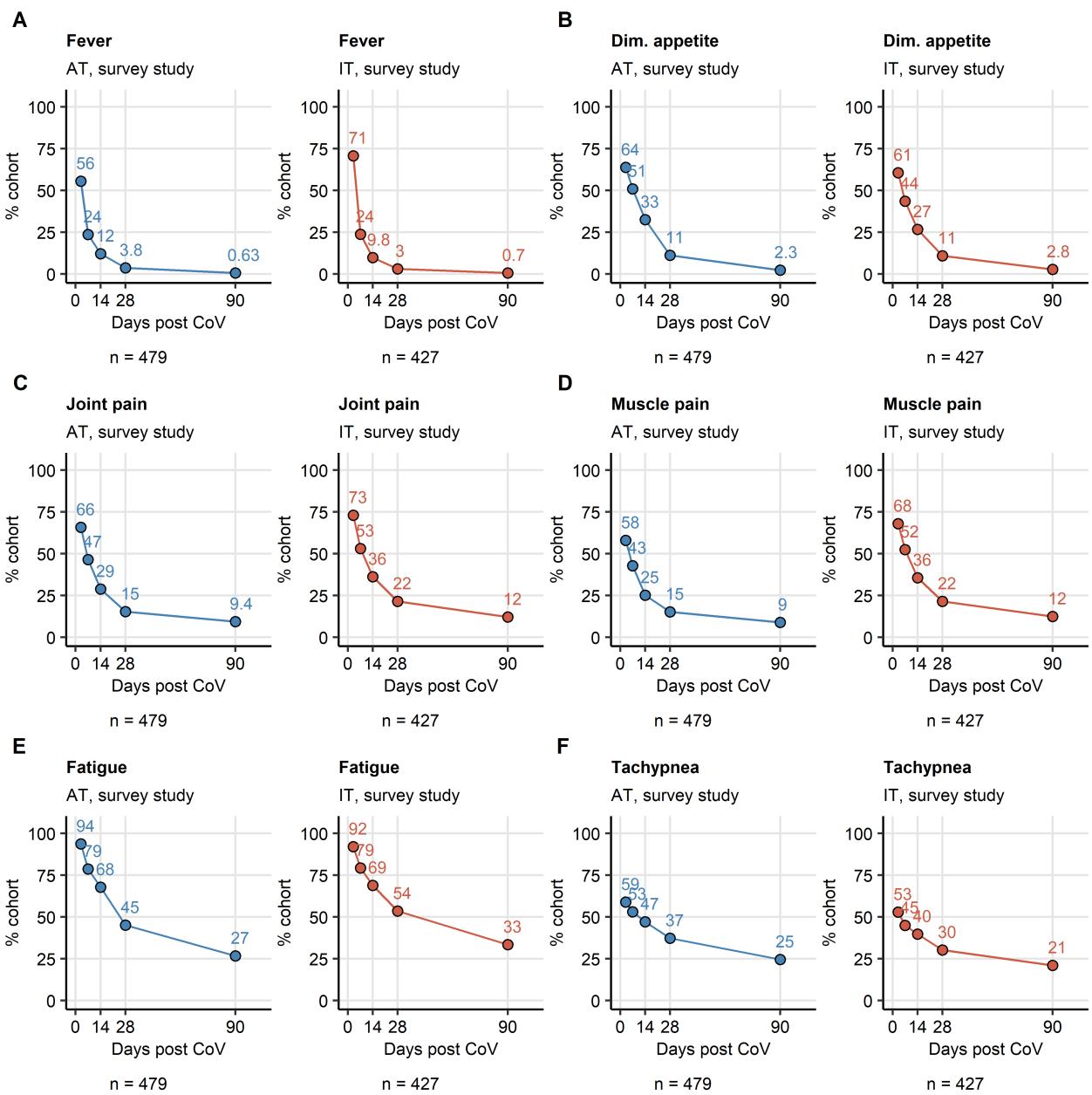
Supplementary Figures



Supplementary Figure S1. Frequency of COVID-19 symptoms in the survey study.

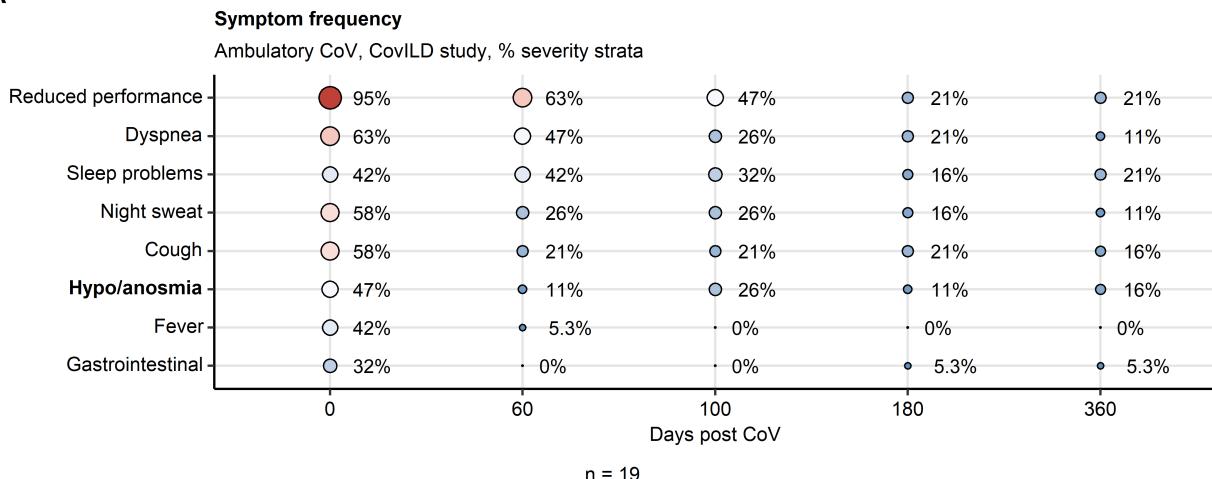
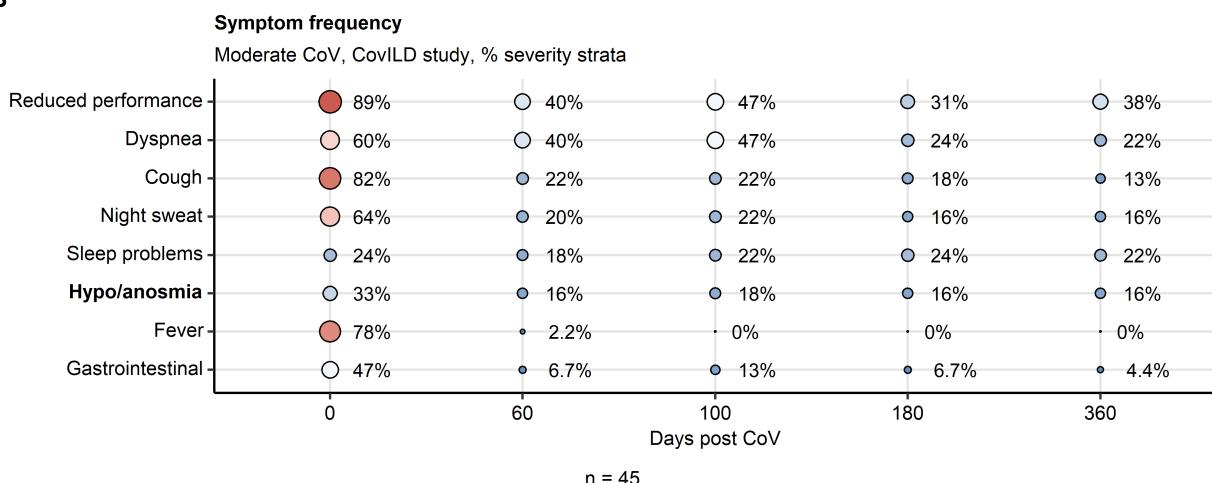
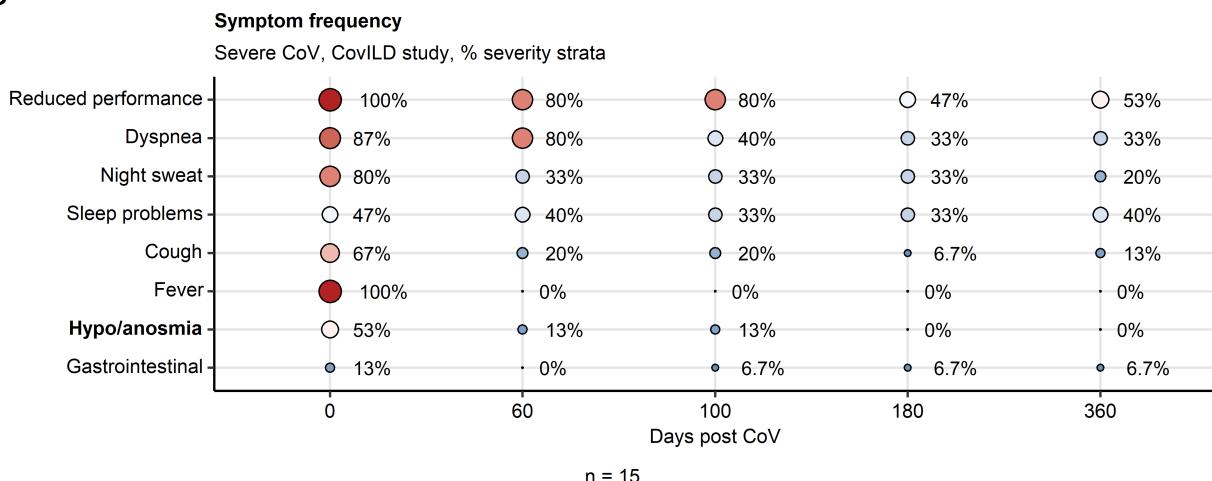
Frequency of symptoms in first 14 days, \geq 28 days and \geq 3 months after clinical onset of COVID-19 in the Austria (AT, A) and Italy (IT, B) survey study cohorts expressed as percentages of the cohort. Point size and color represents the percentage. Numbers of complete observations are indicated under the plot.

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



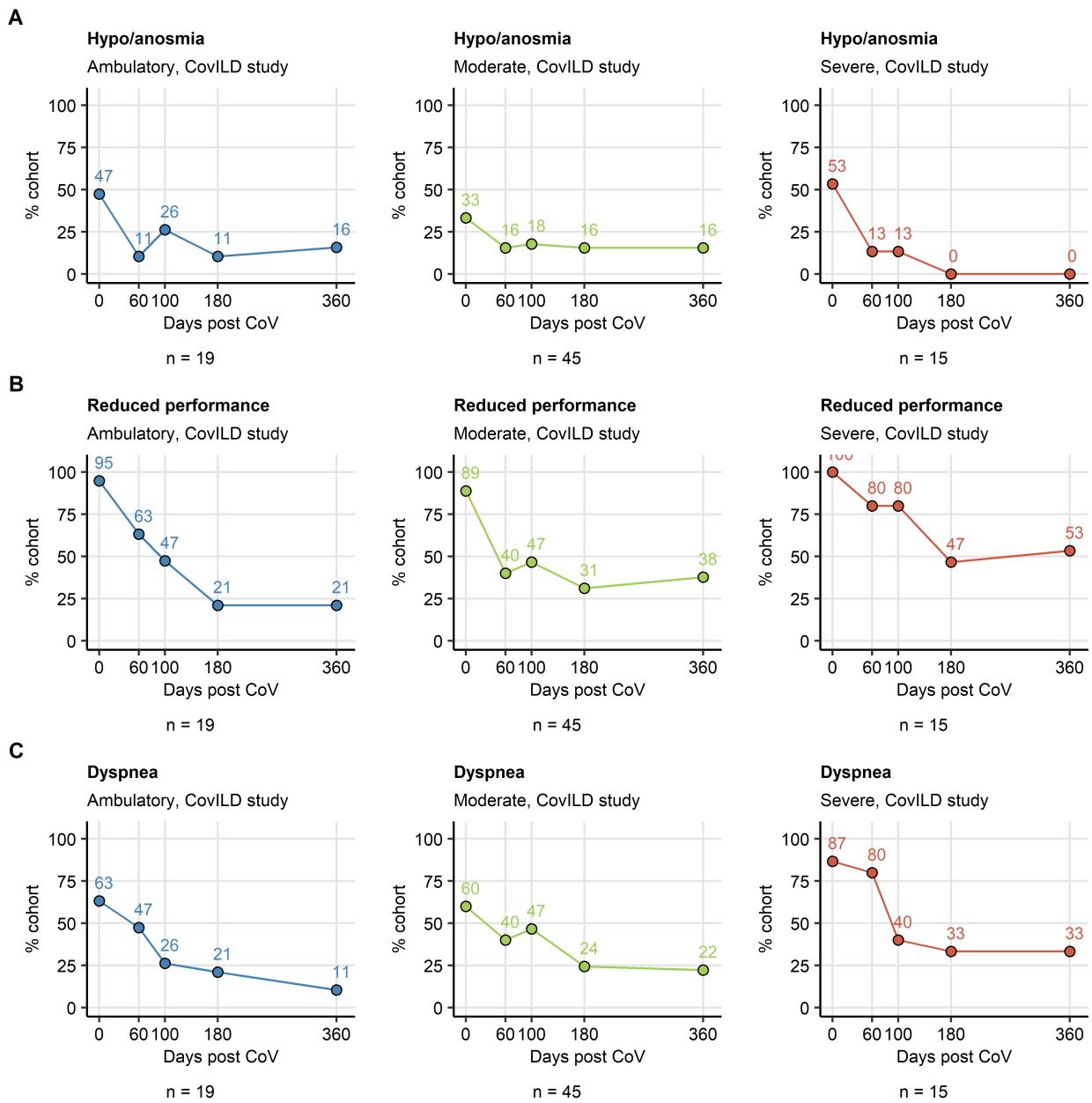
Supplementary Figure S2. Kinetic of recovery from leading acute COVID-19 symptoms in the survey study.

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) survey study cohorts at particular time points after clinical onset. Numbers of complete observations are indicated under the plots.

A**B****C**

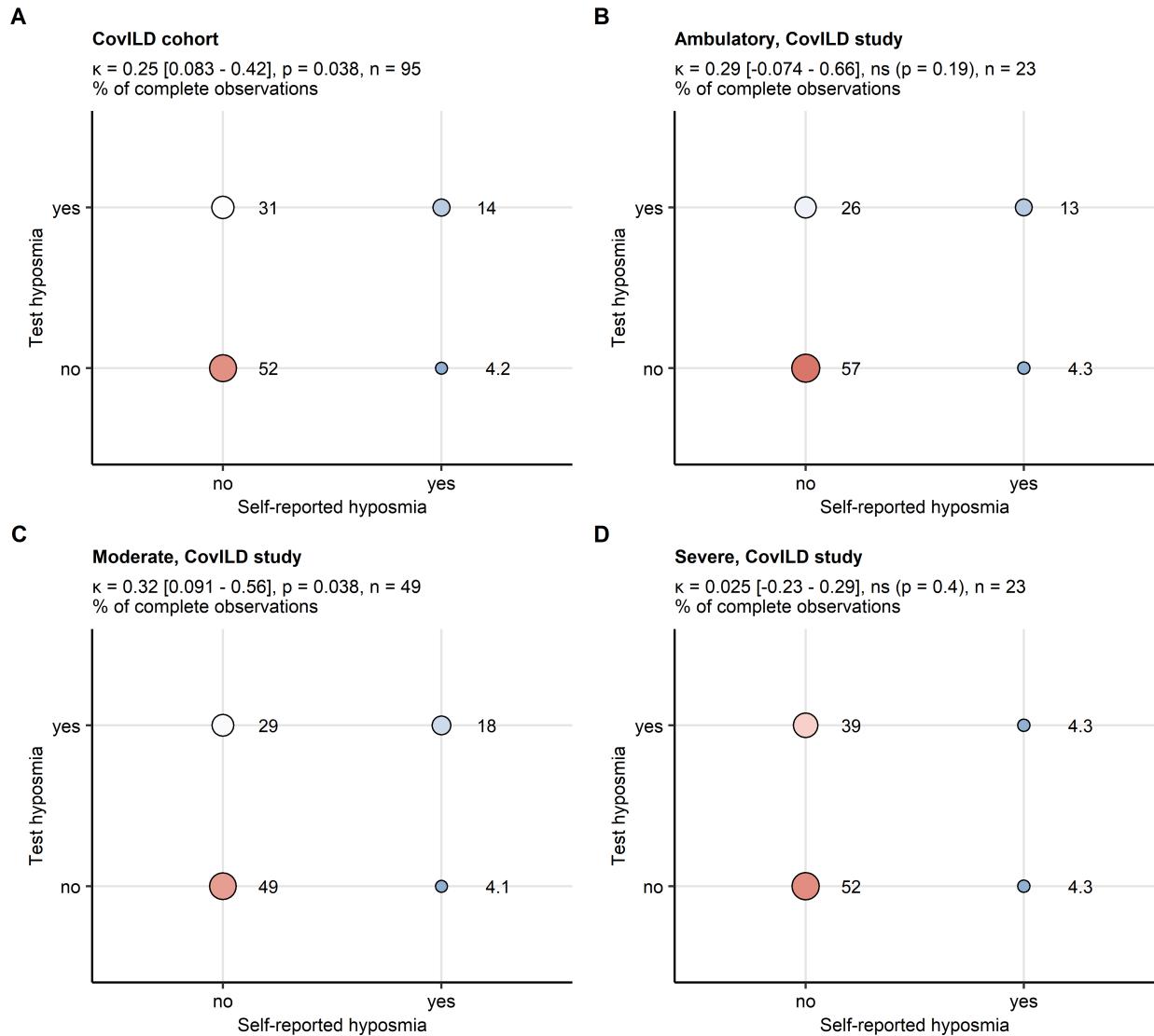
Supplementary Figure S3. Symptom frequency in ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

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 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.



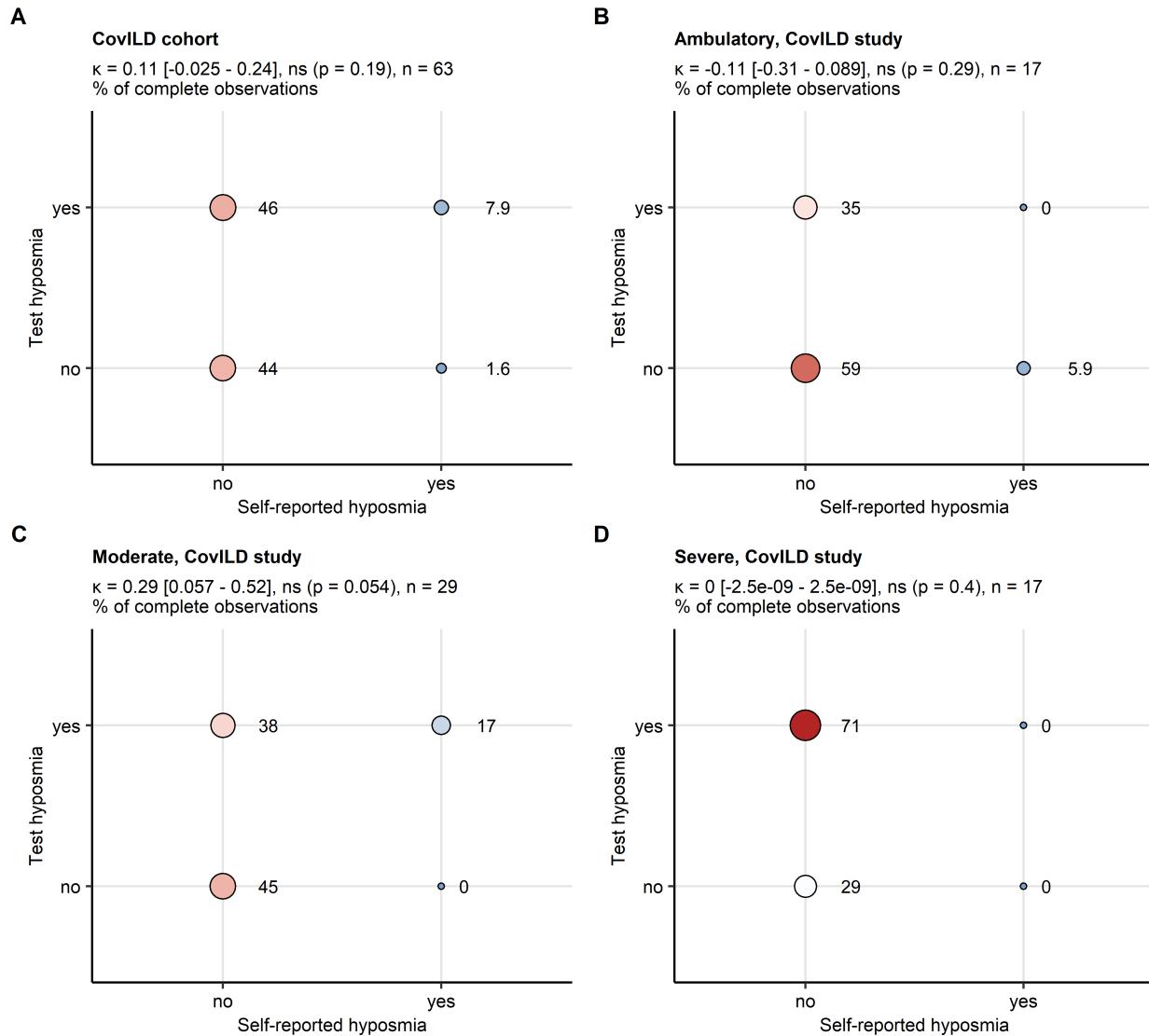
Supplementary Figure S4. Kinetic of recovery from smell disorders, reduced performance and dyspnea in ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

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 during acute COVID-19 and at the 60-, 100-, 180- and 360-day follow-ups. Numbers of complete observations are indicated under the plots.



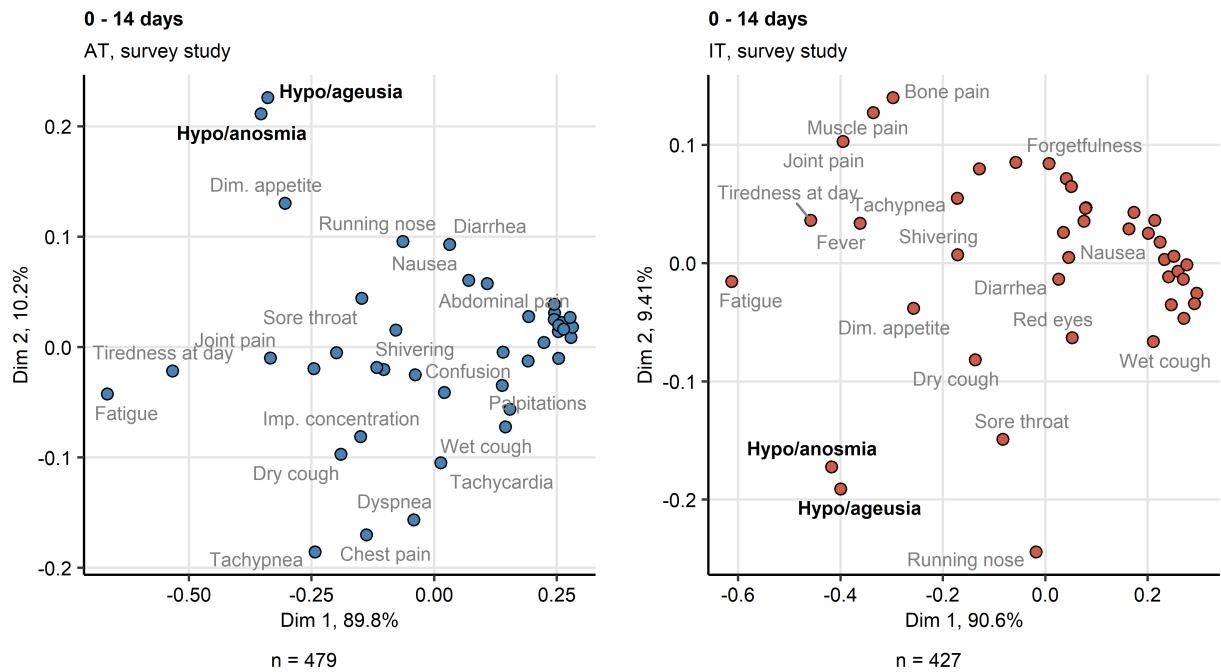
Supplementary Figure S5. Rates of self-reported hyposmia and hyposmia in sniffing stick test at 3-month post COVID-19 follow-up in the ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

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 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.

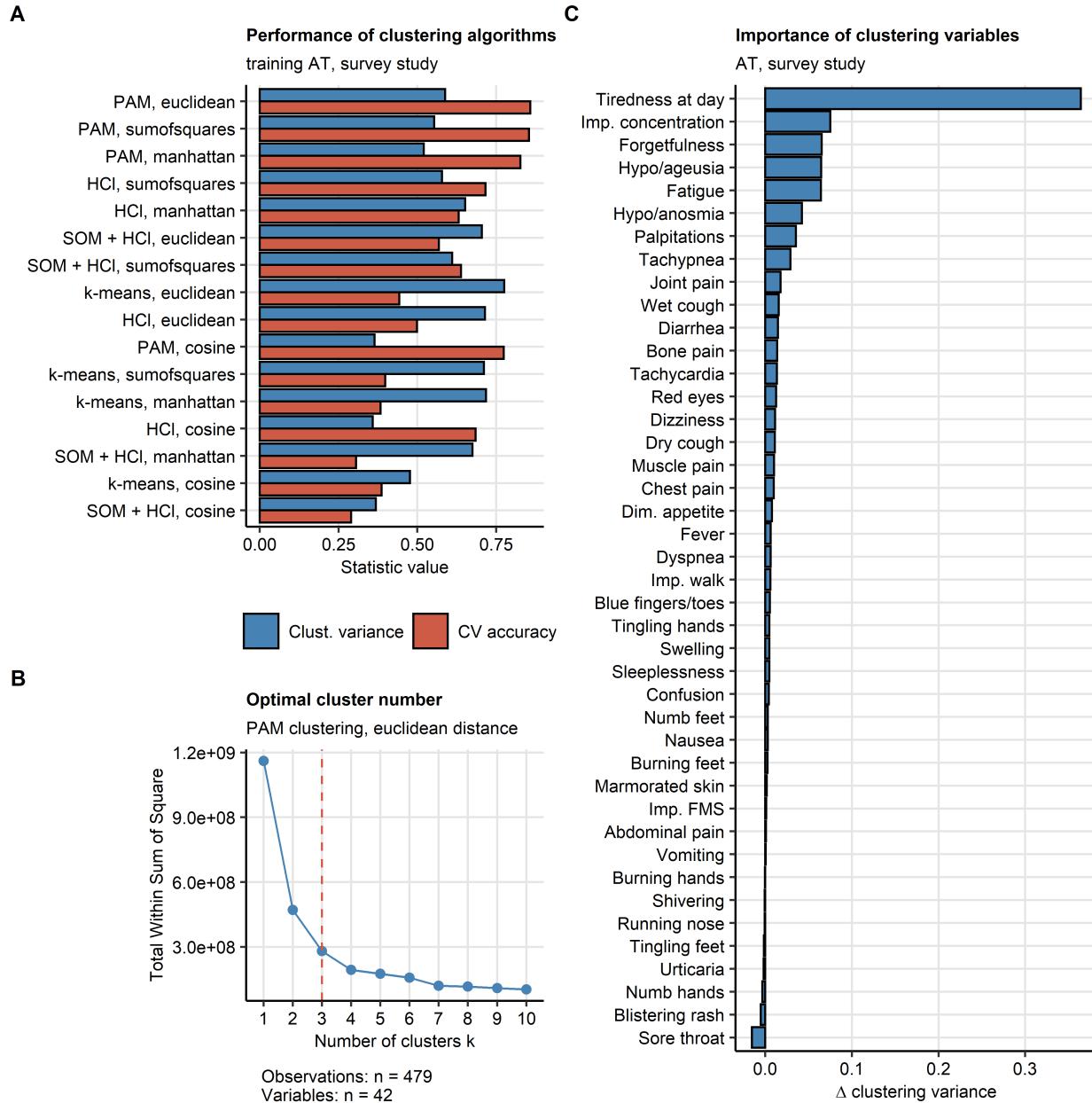


Supplementary Figure S6. Rates of self-reported hyposmia and hyposmia in sniffing stick test at 1-year post COVID-19 follow-up in the ambulatory, moderate and severe COVID-19 subsets of the CovILD study.

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 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 survey study. Symptom data for acute COVID-19 (first 14 days after clinical onset) in the Austria (AT) and Italy (IT) survey study 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 survey study.

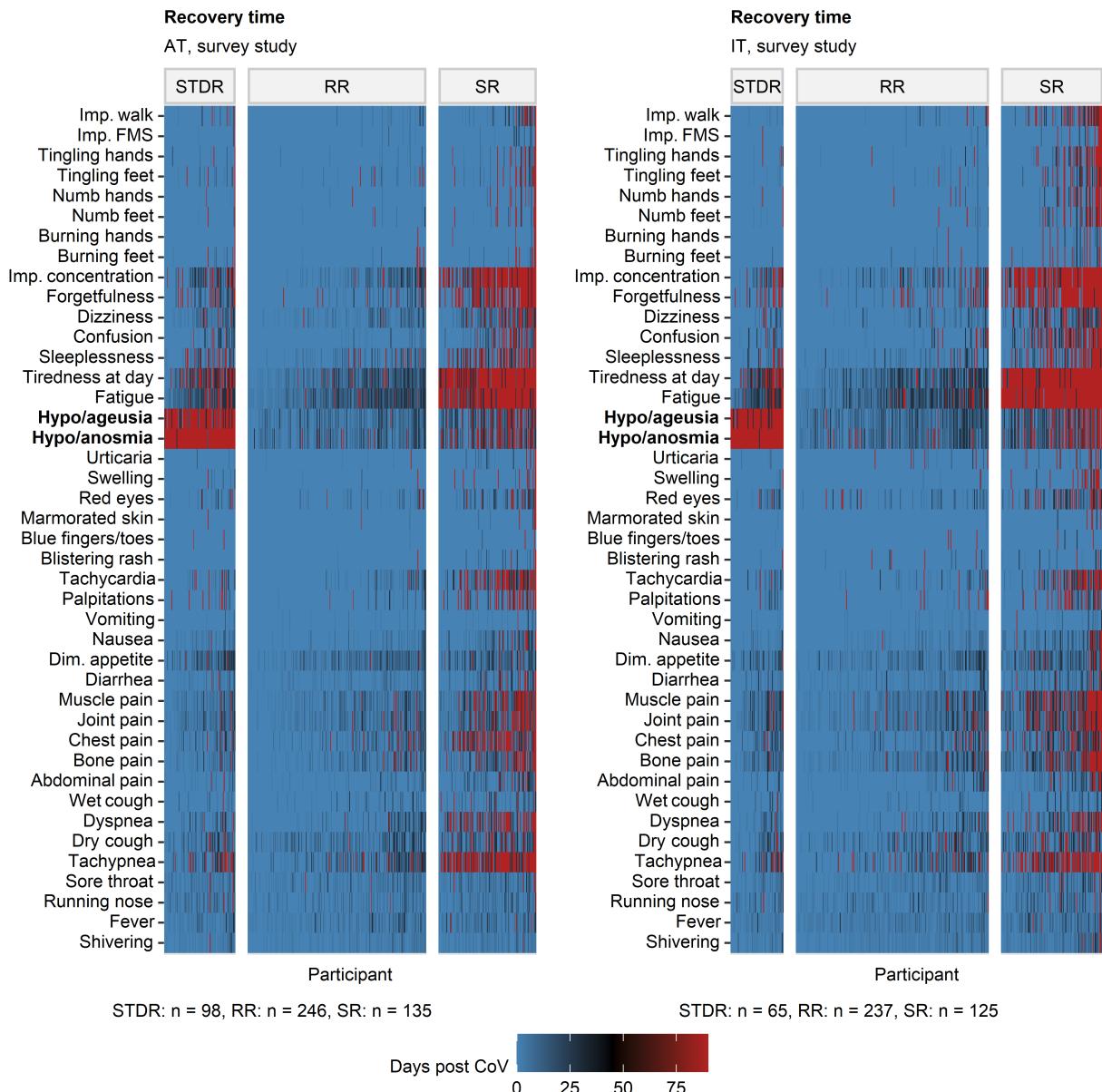
Individuals of the training Austria (AT) study survey cohort were subjected to clustering in respect to symptom-specific recovery times with the PAM (partitioning around medoids) algorithm and Euclidean distance measure.

(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

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 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.

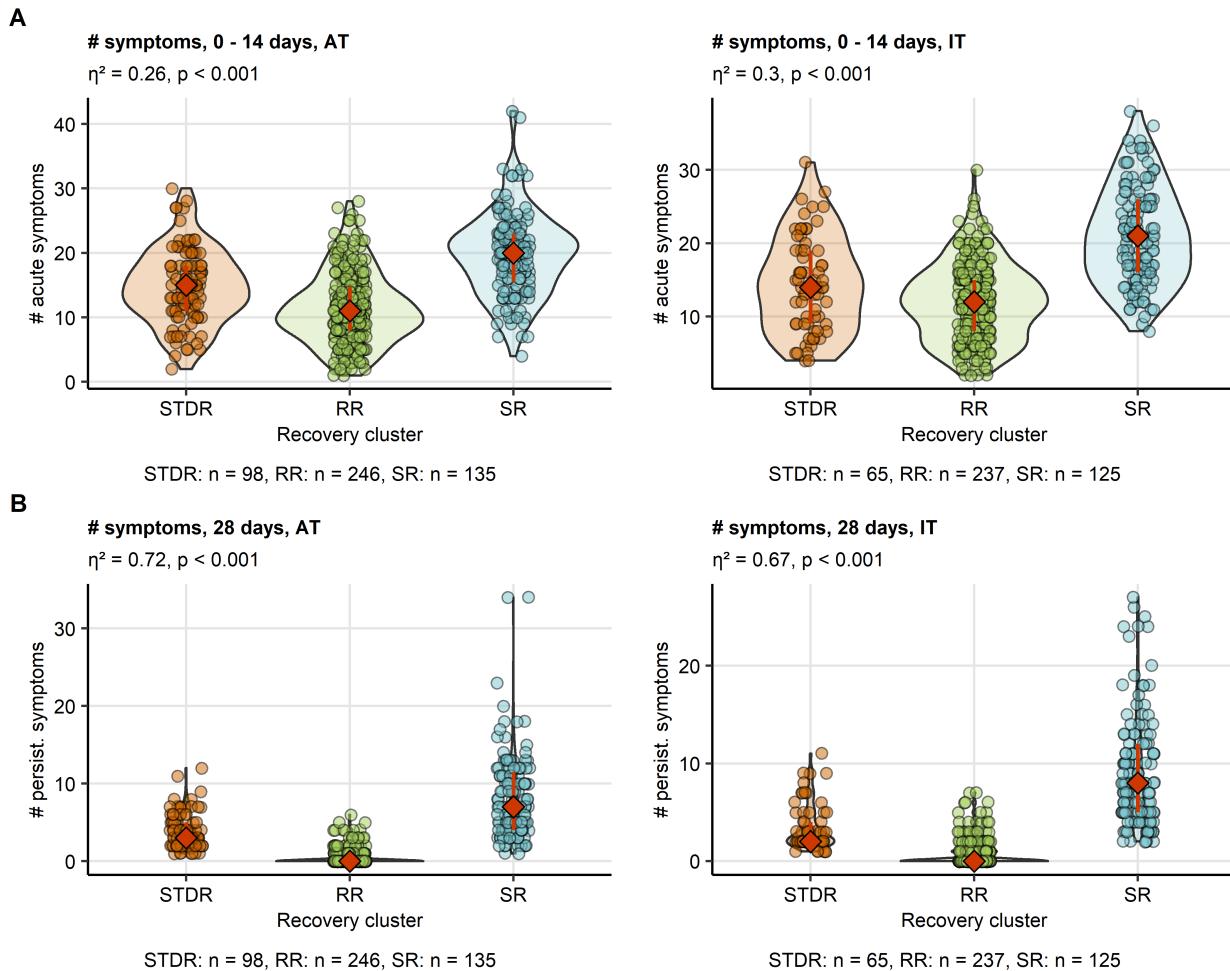


Supplementary Figure S9. Clustering of ambulatory COVID-19 individuals by symptom-specific recovery times.

Individuals of the training Austria (AT) survey study cohort were subjected to clustering in respect to symptom-specific recovery times with the PAM (partitioning around medoids) algorithm and Euclidean distance measure (Supplementary Figure S8). Cluster assignment in the test Italy (IT) survey cohort was done with k-NN label propagation algorithm. Recovery times for particular COVID-19 symptoms in the COVID-19 recovery clusters are presented as

heat maps. Numbers of individuals assigned to the recovery clusters are indicated under the plots.

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



Supplementary Figure S10. Numbers of COVID-19 symptoms in the survey study recovery clusters.

The smell and taste disorder recovery (STDR), rapid recovery (RR) and slow recovery (SR) clusters of the survey study participants were identified by semi supervised PAM clustering (partitioning around medoids, Euclidean distance, training cohort: Austria [AT], test cohort: Italy [IT]). Differences in numbers of symptoms in the first 14 days (A) and at ≥ 28 days (B) after clinical onset 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.

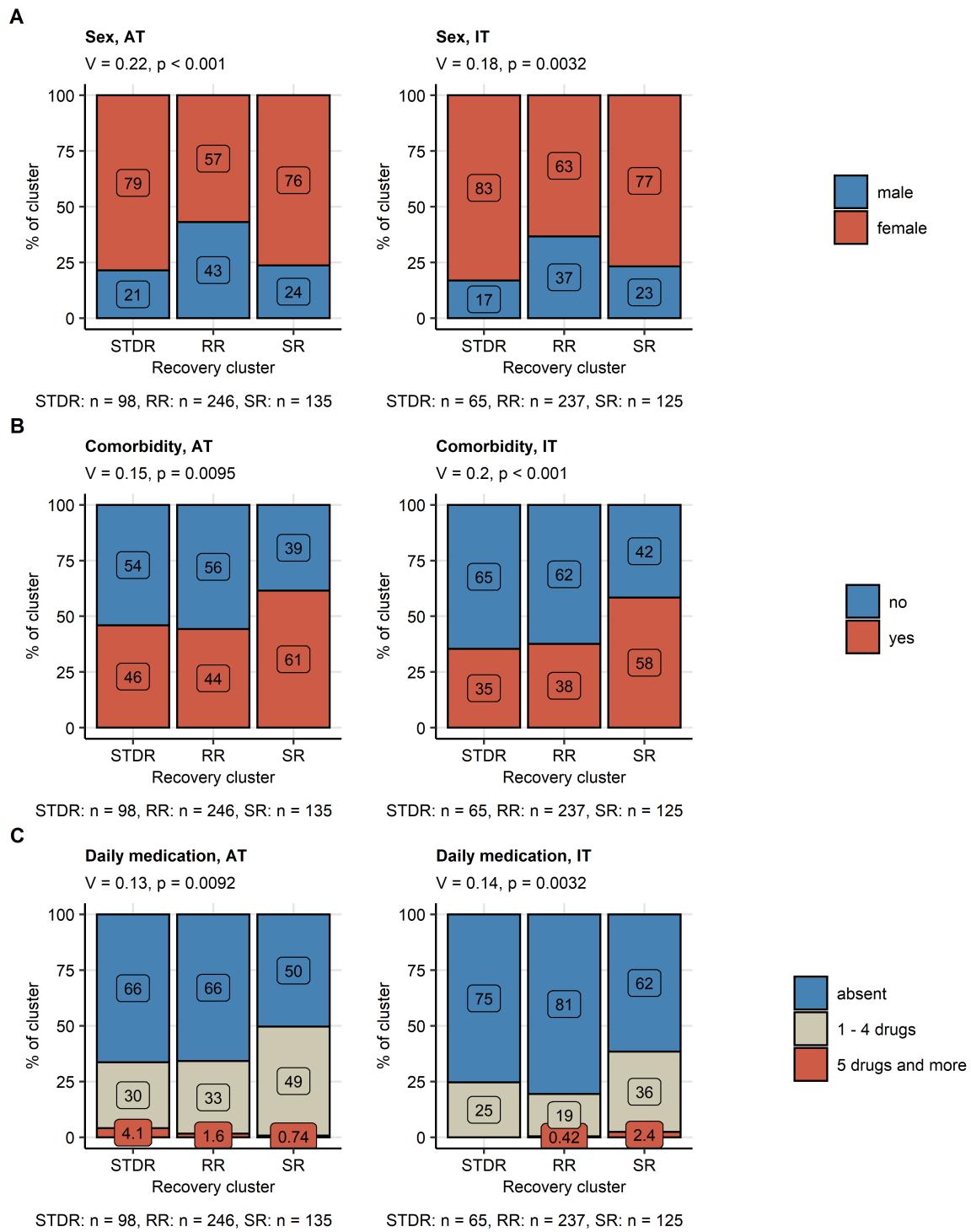


Figure 11. COVID-19 recovery clusters differ in sex distribution, comorbidity and daily medication rates.

The smell and taste disorder recovery (STDR), rapid recovery (RR) and slow recovery (SR) clusters of the survey study participants were identified by semi supervised PAM clustering (partitioning around medoids, Euclidean distance, training cohort: Austria [AT], test cohort: Italy [IT]). Differences in sex distribution (A), frequency of comorbidity (B) and daily medication (C) between the recovery clusters were assessed by χ^2 test with Cramer V effect size statistic. P values were corrected for multiple testing with Benjamini-Hochberg method. The frequencies are presented as bar plots. Effect size statistics and p values are indicated in the plot caption. Numbers of complete observations are displayed under the plots.

References

1. 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](https://doi.org/10.1183/13993003.03481-2020)
2. 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](https://doi.org/10.7554/ELIFE.72500)
3. 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](https://doi.org/10.3389/FMED.2022.792881)
4. 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](https://doi.org/10.1093/CID/CIAB978)
5. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, Schneider A, Brähler E. A 4-item 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](https://doi.org/10.1016/j.jad.2009.06.019)
6. 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](https://doi.org/10.1026/0012-1924.50.4.171)
7. 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](https://doi.org/10.1111/ene.14803)
8. 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](https://doi.org/10.1007/S00405-006-0173-0)

9. 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](https://doi.org/10.21105/joss.01686)
10. Wickham Hadley. *ggplot2: Elegant Graphics for Data Analysis*. 1st ed. New York: Springer-Verlag (2016). <https://ggplot2.tidyverse.org>
11. Wilke CO. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. 1st ed. Sebastopol: O'Reilly Media (2019).
12. 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](https://doi.org/10.1037/h0028106)
13. 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](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)
14. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2021) <https://cran.r-project.org/package=rstatix>
15. Meyer D, Zeileis A, Hornik K. vcd: Visualizing Categorical Data. (2021) <https://cran.r-project.org/web/packages/vcd/index.html>
16. 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](https://doi.org/10.18637/jss.v067.i01)
17. 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](https://doi.org/10.18637/jss.v082.i13)
18. Box GE, Hunter SJ, Hunter WG. Statistics for experimenters: an introduction to design, data analysis, and model building. (2005)
19. Boriah S, Chandola V, Kumar V. Similarity measures for categorical data: A comparative evaluation. *Society for industrial and applied mathematics - 8th SIAM international conference on data mining 2008, proceedings in applied mathematics* 130. (2008). p. 243–254 doi: [10.1137/1.9781611972788.22](https://doi.org/10.1137/1.9781611972788.22)
20. 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>
21. 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](https://doi.org/10.1145/170036.170072)
22. 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](https://doi.org/10.18637/JSS.V014.I15)

23. Drost H-G. Philentropy: Information Theory and Distance Quantification with R. *Journal of Open Source Software* (2018) 3:765. doi: [10.21105/joss.00765](https://doi.org/10.21105/joss.00765)
24. Schubert E, Rousseeuw PJ. Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*. Springer (2019). p. 171–187 doi: [10.1007/978-3-030-32047-8_16](https://doi.org/10.1007/978-3-030-32047-8_16)
25. Lange T, Roth V, Braun ML, Buhmann JM. Stability-Based Validation of Clustering Solutions. *Neural Computation* (2004) 16:1299–1323. doi: [10.1162/089976604773717621](https://doi.org/10.1162/089976604773717621)
26. 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](https://doi.org/10.3923/JSE.2014.14.22)
27. 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>