Frequency of Abnormally Low Neuropsychological Scores in Post-COVID-19 Syndrome: the Geneva COVID-COG Cohort

P. Voruz[1,2,3], I. Jacot de Alcântara[1,2], A. Nuber-Champier[1], A. Cionca[1], G. Allali[2,3,4], L. Benzakour[3,5], P.H. Lalive[2,3], K.-O. Lövblad[3,6], O. Braillard[7], M. Nehme[7], M. Coen[8], J. Serratrice[8], J.-L. Reny[8], J. Pugin[3,9], I. Guessous[3,7], R. Ptak[3,10], B.N. Landis[3,11], F. Assal[2,3,†], J.A. Péron[1,2,†,*]

[1]Clinical and Experimental Neuropsychology Laboratory, Faculty of Psychology, University of Geneva, Geneva, Switzerland
[2]Department of Clinical Neurosciences, Neurology Department, Geneva University Hospitals, Geneva, Switzerland
[3]Faculty of Medicine, University of Geneva, Geneva, Switzerland
[4]Leenaards Memory Center, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[5]Psychiatry Department, Geneva University Hospitals, Geneva, Switzerland
[6]Diagnostic and Interventional Neuroradiology Department, Geneva University Hospitals, Geneva, Switzerland
[7]Division and Department of Primary Care, Geneva University Hospitals, Geneva, Switzerland
[8]Division of General Internal Medicine, Department of Medicine, Geneva University Hospitals and Geneva University, Geneva, Switzerland
[9]Intensive Care Department, Geneva University Hospitals, Geneva, Switzerland
[10]Neurorehabilitation Department, Geneva University Hospitals, Geneva, Switzerland
[11]Rhinology-Olfactology Unit, Otorhinolaryngology Department, Geneva University Hospitals, Geneva Switzerland

*Corresponding author at: Faculté de Psychologie et des Sciences de l'Education, 40 bd du Pont d'Arve, 1205 Geneva, Switzerland.
Tel.: +41-22-379-94-55; Fax: 0041 22 379 93 59.
E-mail address: julie.peron@unige.ch (J.A. Péron).
†These authors contributed equally to this work.

Abstract

Objective: Several studies have reported poor long-term neuropsychological performances in patients following severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, but none has yet considered the effect of administering multiple intercorrelated neuropsychological tests and assessed the frequency of cognitive deficits in a normative population. Our aim was therefore to assess the presence of cumulative neuropsychological deficits in an actual post-coronavirus disease of 2019 (COVID-19) comparison group versus one simulated using Monte-Carlo methods.
Method: Validated neuropsychological Monte-Carlo simulation methods were applied to scores from a battery of neuropsychological tests (memory, executive, attentional, perceptual, logical reasoning, language, and ideomotor praxis) administered to 121 patients who had had mild, moderate, or severe COVID-19 (mean age: 56.70 years; 32% women), 222 ± 43 days post-infection. The cumulative percentages of the three severity subgroups were compared with the results of a false discovery rate-corrected probability analysis based on normative data.
Results: The cumulative percentages of deficits in memory and executive functions among the severe and moderate patients were significantly higher than those estimated for the normative population. Moderate patients also had significantly more deficits in perception and logical reasoning. In contrast, the mild group did not have significantly more cumulative deficits.
Conclusions: Moderate and severe forms of COVID-19 cause greater long-term neuropsychological deficits than those that would be found in a normative population, reinforcing the hypothesis of long-term effects of SARS-CoV-2 on cognitive function, independent of the severity of the initial infection.

Keywords: Neuropsychology; Post-COVID-19 condition; Simulation; Memory; Executive functions

Little is known about the frequency of long-term neuropsychological deficits following severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. To date, most studies have assessed the frequency of neuropsychological deficits using global cognitive efficiency tests such as the Montreal Cognitive Assessment (MoCA), Mini-Mental State Examination (MMSE), and Telephone Interview for *Cognitive* Status (for review see Daroische, Hemminghyth, Eilertsen, Breitve, & Chwiszczuk, 2021). Few studies have assessed this frequency using neuropsychological tests that probe specific cognitive functions (Almeria, Cejudo, Sotoca, Deus, & Krupinski, 2020; Jaywant et al., 2021; Krishnan, Miller, Reiter, & Bonner-Jackson, 2022). Adopting a conservative methodology (5th percentile cutoff), Almeria and colleagues (2020) reported impaired (5.70%–11.40%) scores on all assessed functions 10–35 days post-discharge on average after a SARS-CoV-2 infection. Group comparisons on raw scores revealed significantly lower scores for patients who had been in an intensive care unit (ICU) than for patients who had been hospitalized but not in ICU, as well as for patients who had received oxygen therapy, versus those who had not (Almeria et al., 2020). When Jaywant and colleagues (2021) used a combination of conservative and nonconservative (borderline scores below 16th percentile) methodologies to assess the frequency of memory, executive, and attention deficits in 57 patients 43.2 days after initial admission, they observed deficits in all functions (46% of the sample in the case of divided attention). Almeria and colleagues (2020) and Jaywant and colleagues (2021) highlighted the short-term frequency of deficits (<3-month post-infection), whereas Krishnan and colleagues (2022) assessed 20 mild-to-moderate patients 5.5 months post-infection and highlighted a frequency of deficits ranging from 5% to 52% for memory and executive function tests, suggesting a potentially high frequency of long-term neuropsychological deficits in patients with mild infection. However, these studies had two main methodological biases. First, there was a lack of analysis to determine whether the neuropsychological deficits were greater than would be expected in a normative population, depending on the number of tests performed (the higher the number of neuropsychological tests, the greater the probability of observing a subnormal neuropsychological score, even in a normative population; Crawford, Garthwaite, & Gault, 2007). Over the past two decades, studies among populations of neurotypical participants without any cognitive complaints have shown that the application of frequency calculation methodologies, be they conservative or nonconservative, can lead to the discovery of a high percentage of deficits in the normative population (for a review, see Binder, Iverson, & Brooks, 2009). Furthermore, the risk of intercollinearity (i.e., two or more variables being completely correlated) of the neuropsychological tests was not considered in these studies, even though it could potentially indicate an accumulation of deficits in the various cognitive functions being assessed. Running simulation analyses of neuropsychological data, as proposed by Crawford and colleagues (2007), would therefore allow authors to determine whether a sample's accumulated deficits are higher than would be expected in a normative population or in a clinical subgroup of patients. Second, a lack of analysis with dedicated tools for evaluating the presence of symptoms to be considered with care and the validity of responses may have led to a superficially high frequency of neuropsychological deficits (Binder et al., 2009).

In this context, the present study had three main objectives: (a) investigate whether SARS-CoV-2 infection causes long-term (6–9 months after acute phase) neuropsychological deficits, controlling for the validity of symptoms; (b) identify the nature of the affected cognitive domains using a comprehensive neuropsychological assessment to probe multiple cognitive domains; and (c) determine whether the cumulative percentage of neuropsychological deficits as a function of severity of acute infection is higher than would be expected in a normative population, by considering possible intercorrelations between the neuropsychological tests used. Patients were divided into three subgroups according to the respiratory severity (severe, moderate, and mild) of their disease in the acute phase, and their cumulative percentages were then compared with scores estimated for a normative population, based on normative data. Previous studies had highlighted impaired neuropsychological performances (raw scores; Almeria et al., 2020; Jaywant et al., 2021) and a higher frequency of deficits in mild-to-moderate patients (Krishnan et al., 2022). We hypothesized that coronavirus disease of 2019 (COVID-19) causes long-term neuropsychological deficits that continue to affect patients' functioning 6–9 months post-infection. More specifically, we expected to observe a higher cumulative percentage (frequency) of abnormally low scores on memory and executive functions among patients who had moderate-to-severe acute infections (hospitalized) than that estimated for the normative population using the Monte-Carlo simulation method developed by Crawford and colleagues (2007), validated by Schretlen, Testa, Winicki, Pearlson, and Gordon (2008), and adapted for our study.

Methods

Participants

Of all the patients who had an infection confirmed by positive polymerase chain reaction (PCR) results from a nasopharyngeal swab and/or by positive serological results between March 2020 and May 2021 at Geneva University Hospitals (HUG) (∼4,000 patients), ∼300 patients met the inclusion criteria (see below), and of these a total of 121 agreed to participate in the study 222.46 ± 42.93 days after their infection. According to the severity of their infection in the acute phase, these patients were

Table 1. Sociodemographic data and relevant medical history

	Total sample n = 121	Mild subgroup n = 49	Moderate subgroup n = 48	Severe subgroup n = 24
Mean age in years (± SD)	56.69 (± 10.41)	54.86 (± 8.78)	55.85 (± 10.40)	62.08 (± 12.03)
[range]	[35–78]	[35–71]	[36–75]	[38–78]
Mean education level [1–3][a] (± SD)	2.65 (± 0.54)	2.76 (± 0.43)	2.63 (± 0.61)	2.50 (± 0.59)
Sex (% woman)	32.23	34.69	35.42	20.83
Mean days of hospitalization (± SD)	—	—	11.89 (± 11.89)	40.13 (± 32.07)
Mean days between infection and assessment (± SD)	222.46 (± 42.93)	215.98 (± 35.79)	223.65 (± 50.50)	233.04 (± 38.69)
Diabetes (%)	8.26	2.04	8.33	20.83
Smoking (%)	6.61	12.24	2.08	4.20
History of respiratory disorders (%)	14.88	12.24	12.50	25.00
History of cardiovascular disorders (%)	16.53	12.24	16.67	25.00
History of neurological disorders (%)	0	0	0	0
History of psychiatric disorders (%)	3.31	4.08	2.08	4.20
History of cancer (%)	0	0	0	0
History of severe immunosuppression (%)	0	0	0	0
History of developmental disorders (%)	0	0	0	0
Chronic renal failure (%)	1.65	0	0	8.30
Sleep apnea syndrome (%)	14.88	8.16	14.58	29.20

Note. Mild: patients not hospitalized for SARS-CoV-2 infection; Moderate: patients hospitalized without mechanical ventilation for SARS-CoV-2 infection; *ns*: not significant; Severe: patients hospitalized in intensive care with mechanical ventilation for SARS-CoV-2 infection; *SD*: standard deviation.
[a]Level 1 is equivalent to the compulsory Swiss scholarship (<11 years of study); level 2 is equivalent to a vocational diploma (11–12 years of study); and level 3 is equivalent to Matura level and higher education (>12 years of study).

divided into three subgroups: mild (not hospitalized; $n = 49$), moderate (hospitalization without mechanical ventilation; $n = 48$), or severe (ICU and mechanical ventilation; $n = 24$; see Table 1).

During the inclusion phase, the mild and moderate groups were matched to patients in the severe group so that the three groups would be comparable on median age, sociocultural level, language (all were French-speaking Swiss citizens or residents of the French part of Switzerland), and clinical variables (see Table 1). Participants were recruited via the CoviCare program (Nehme et al., 2021), which follows patients with post-COVID symptoms in Geneva, Switzerland (MN, OB, and IG), as well as from the registers of another study (LB; Benzakour et al., 2021). For each patient, we carried out a medical file review, followed by a telephone call inviting the patient to take part in the study if all the eligibility criteria were met. Exclusion criteria were a history of neurological issues, psychiatric disorders (two of the included participants had had an episode of depression >10 years before their SARS-CoV-2 infection), cancer and neurodevelopmental pathologies, pregnancy, and age above 80 years. The frequency of the different clinical symptoms experienced by patients during the acute phase is indicated in Table 2.

General Procedure and Ethics

After being given a full description of the study, participants provided their written informed consent. The study was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the cantonal ethics committee of Geneva (CER-02186).

Neuropsychological Assessment

A comprehensive neuropsychological battery solely containing tests validated in a French-speaking population was administered in French (all patients were fluent in written and spoken French). This battery included a series of tests and questionnaires that assessed most of the domains of cognition (for testing details, see Table 3). The tests were administered by certified clinical psychologists (mean duration: ∼180 min). A neurological assessment was carried out by two board-certified neurologists (FA and GA; see Voruz et al., 2022a).

Measure of Symptom Validity

We used the BRIEF-A (Abeare et al., 2021) and WAIS-IV digit-span measures (Kanser, Rapport, Hanks, & Patrick, 2021) to measure the validity of the patients' performance, the validity (i.e., congruence) of their symptoms, and the presence of any symptoms to be considered with care. The default validity check of the BRIEF-A (negativity scale ≥ 6) yielded good-to-excellent

Table 2. Percentages of self-reported symptoms in the acute phase as a function of severity of respiratory symptoms

	Total sample $N = 121$	Mild subgroup $n = 49$	Moderate subgroup $n = 48$	Severe subgroup $n = 24$
Runny nose	25.47%	32.65%	18.75%	25.00%
Sore throat	16.51%	22.45%	14.58%	12.50%
Muscle pain	49.50%	71.43%	39.58%	37.50%
Loss of sense of smell	39.20%	55.10%	50.00%	12.50%
Taste disorder	37.85%	48.98%	47.92%	16.67%
Dry cough	54.49%	53.06%	56.25%	54.17%
Productive cough	6.16%	12.24%	6.25%	0.00%
Fever	0.00%	0.00%	0.00%	0.00%
Digestive symptoms	67.64%	59.18%	72.92%	70.83%
Fatigue	33.11%	32.65%	45.83%	20.83%
Difficulty breathing	76.50%	83.67%	83.33%	62.50%
Chest pain	45.59%	34.69%	52.08%	50.00%
Headache	23.44%	24.49%	25.00%	20.83%
Somnolence	26.12%	38.78%	27.08%	12.50%
Nonrestorative sleep	55.73%	75.51%	54.17%	37.50%
Insomnia	32.41%	32.65%	35.42%	29.17%
Waking up feeling choked or suffocated	17.98%	10.20%	22.92%	20.83%
Snoring	12.46%	6.12%	18.75%	12.50%
Interruption of breathing during sleep	0.69%	0.00%	2.08%	0.00%
Other	5.56%	0.00%	8.33%	8.33%
None	19.30%	20.41%	29.17%	8.33%

Note. Mild: patients not hospitalized for SARS-CoV-2 infection; Moderate: patients hospitalized without mechanical ventilation for SARS-CoV-2 infection; and Severe: patients hospitalized in intensive care with mechanical ventilation for SARS-CoV-2 infection.

Table 3. Domains and functions measured by the neuropsychological tests used in the COVID-COG protocol

Domain	Functions	Names of tests
Perception	Object perception	Incomplete Letters and Object Decision tests from Visual Object and Space Perception battery (Warrington & James, 1991)
	Spatial perception	Number Location and Cube Analysis tests from Visual Object and Space Perception battery (Warrington & James, 1991)
Ideomotor praxis		Moroni praxis battery (Mahieux-Laurent et al., 2009)
Language	Semantic processing: naming and repetition	Semantic image matching, semantic word matching, oral picture naming, word repetition, and nonword repetition from BECLA battery (Macoir et al., 2016)
Executive functions	Inhibition	Stroop task from GREFEX battery (Roussel & Godefroy, 2008)
	Mental flexibility	Trail Making Test from GREFEX battery (Roussel & Godefroy, 2008)
	Verbal fluency	Categorical and Verbal Fluency from GREFEX battery (Roussel & Godefroy, 2008)
	Verbal working memory	Digit Span Backward from WMS-III (Drozdick, Raiford, Wahlstrom, & Weiss, 2018)
	Visuospatial working memory	Backward Corsi test from WAIS-IV (Kessels, Van Zandvoort, Postma, Kappelle, & De Haan, 2000) and Test for Attentional Performance (Zimmermann & Fimm, 2007)
Attention	Phasic alertness; divided and sustained attention; incompatibility	Test for Attentional Performance (Zimmermann & Fimm, 2007).
Memory	Episodic verbal	Grober and Buschke free/cued recall (RL/RI 16) paradigm (Grober & Buschke, 1987)
	Episodic visuospatial	Delayed recall of Rey-Osterrieth Complex Figure test (Meyers & Meyers, 1995)
	Anosognosia for memory dysfunction	Self-appraisal discrepancy (SAD) score for each memory test (Voruz et al., 2022b)
Logical reasoning		Matrix Reasoning and Visual Puzzles subtests from WAIS-IV (Wechsler, 2008)

Note. BECLA = Batterie d'Evaluation Cognitive du Langage (Macoir et al., 2016); GREFEX = Groupe de Réflexion sur l'Evaluation des Fonctions Exécutives; ICU = intensive care unit; WMS-III = Wechsler Memory Scale—Third Edition (Drozdick et al., 2018); Rey Figure = Rey-Osterrieth Complex Figure test; RL/RI 16 = free/cued recall 16 items; *SD* = standard deviation; TAP = Test for Attentional Performance. Version 2.1 (Zimmermann & Fimm, 2007); TMT = Trail Making Test; VOSP = Visual Object and Space Perception battery; WAIS-IV = Wechsler Adult Intelligence Scale–Fourth Edition.

results for all participants, as did our use of the cutoff purposed by Abeare and colleagues (2021) (failure on ≥4 of the individual scales). These good-to-excellent scores on the BRIEF-A, were partially confirmed by the validity measured by the WAIS-IV Digit Span Forward and Backward with a threshold of ≥5 measured in a healthy control population and which simulated

(Kanser et al., 2021). In this case, only one participant had a score above the threshold on the Digit Span Forward, and two patients on the Digit Span Backward. Although they performed well on the BRIEF-A, these patients showed other attentional deficits, as measured with the Test for Attentional Performance Version 2.1 (Zimmermann & Fimm, 2007), suggesting genuine impairment.

Statistical Analyses

For each neuropsychological test, we first compared patients' performances with normative data collected from a reference sample for the validated neuropsychological tools. Raw scores were converted to standardized scores and adjusted for age, education, and gender, using published normative data and in accordance with the guidelines of the Swiss Association of Neuropsychology (Frei et al., 2016; Heaton, Grant, & Matthews, 1991), allowing us to quantify the frequency of each type of impairment.

To determine whether the deficits we observed in our patient cohort were greater than those that would be expected in a normative population, and to compensate for the number of tests, we calculated the frequency of low scores using the Monte-Carlo simulation program (i.e., estimated base rates), with a framework developed by Crawford and colleagues (2007) and validated in the neuropsychology field. This method estimates the baseline rates of low scores according to test intercorrelations (in our case, for the whole sample of patients with SARS-CoV-2). It allowed us to estimate, as recommended by Crawford and colleagues (2007), the percentage of the normative population who would exhibit one or more, two or more, three or more, or four or more abnormally low scores, applying a conservative threshold (< 5th percentile). We carried out six steps. (a) We pooled the scores according to the cognitive functions they were supposed to assess, based on theoretical models and/or test batteries: episodic memory model (Schacter & Tulving, 1994) and The Memory NEo-Structural Inter-Systemic model (MNESIS) model (Eustache, Viard, & Desgranges, 2016) for memory; latent model of executive functions (Miyake et al., 2000) and GREFEX battery (Roussel & Godefroy, 2008) for executive functions; test for attentional performance (TAP) battery (Zimmermann & Fimm, 2007) for attentional abilities; subtests of the VOSP (Warrington & James, 1991) for perceptual abilities; subtests of the WAIS-IV (Wechsler, 2008) for logical reasoning; subtests of the BECLA (Macoir, Gauthier, Jean, & Potvin, 2016) for language; and subtests of a praxis battery (Mahieux-Laurent et al., 2009) for ideomotor praxis; detailed scores provided in Supplementary Index 1. (b) We calculated correlation matrices of raw scores for each function. (c) We fed the results of the correlation matrices into a generic program: PercentAbnormKtests (available at: https://homepages.abdn.ac.uk/j.crawford/pages/dept/PercentAbnormKtests.htm; Crawford et al., 2007) with a conservative threshold (< 5th percentile). (d) We summed scores below the conservative threshold for each patient and each function. (e) We calculated the cumulative percentages of patients in each group (total sample; mild, moderate, and severe subgroups) with at least one lower test score. (f) We compared these cumulative percentages with the estimated scores of the normative population (obtained at Step 3 using binomial distribution probability analyses, as suggested by Crawford and colleagues (2007), with the estimated baseline rate used to specify *p* sample: (i) estimate versus mild group; (ii) estimate versus moderate group; and (iii) estimate versus severe group. Finally, we applied a Benjamini–Hochberg false discovery rate (FDR) correction with a *p* value set at .050 for each group comparison (estimate vs. total ere) on each of the functions we assessed (memory, executive, attentional, perceptual, logical reasoning, language, and ideomotor praxis).

Results

Perceptual Functions

The total sample and the normative population only differed significantly on four or more abnormally low scores [total sample has 0.82% more abnormally low scores than the estimated score for the normative population (+ 0.82%, $p < .001$)]. All other comparisons were nonsignificant ($ps > .05$; Table 4).

For the moderate subgroup, significant differences were found with the normative population on two or more (+ 8.09%, $p < .001$), three or more (+ 1.86%, $p < .001$), and four and more (+ 2.07%, $p < .001$) abnormally low scores. No other significant differences were found (all $ps > .05$).

For the mild and severe subgroups, all cumulative percentages were nonsignificant, compared with the normative population (all $ps > .05$).

Language and Ideomotor Praxis

Analysis revealed no significant differences between the normative population and either the total sample or the mild, moderate, and severe subgroups (all $ps > .05$; Table 4).

Table 4. Comparisons between cumulative percentages of abnormally low neuropsychological scores of post-COVID-19 groups (total sample; mild, moderate, and severe subgroups) 6–9 months post-infection and those estimated for a normative population

		Estimated cumulative percentages of deficits for normative population (< 5th percentile)	Frequency (%) of scores—conservative methodology (< 5th percentile)			
			Total sample $N = 121$	Mild subgroup $n = 49$	Moderate subgroup $n = 48$	Severe subgroup $n = 24$
Perception (4 tests)	0 low scores	82.53%	80.99%	83.67%	77.08%	83.33%
	≥ 1 low scores	17.47%	19.08%	16.32%	22.92%	16.67%
	≥ 2 low scores	2.33%	4.96%	0%	10.42%**	0%
	≥ 3 low scores	0.22%	0.83%		2.08%**	0%
	4 low scores	0.01%	0.83%**		2.08%**	
Ideomotor praxis (3 tests)	0 low scores	85.97%	90.08%	89.80%	91.67%	87.50%
	≥ 1 low scores	14.03%	9.92%	10.20%	8.33%	12.50%
	≥ 2 low scores	0.94%	0%	0%	0%	0%
	3 low scores	0.02%				
Language (5 tests)	0 low scores	79.69%	82.64%	85.71%	81.71%	79.17%
	≥ 1 low score	20.31%	17.36%	14.28%	14.58%	20.83%
	≥ 2 low score	4.04%	3.31%	4.08%	4.16%	0%
	≥ 3 low scores	0.65%	0%	0%	0%	
	≥ 4 low scores	0.08%				
	5 low scores	0.00%				
Executive functions (11 tests)	0 low scores	68.51%	54.55%	69.39%	41.67%	50%
	≥ 1 low score	31.49%	45.45%**	30.61%	58.33%**	50%**
	≥ 2 low score	12.71%	25.62%**	16.33%	33.33%**	29.17%**
	≥ 3 low scores	5.87%	11.57%**	6.12%	16.67%**	12.5%
	≥ 4 low scores	2.82%	6.61%**	2.04%	12.50%**	4.16%
	≥ 5 low scores	1.32%	1.65%	0%	4.17%	0%
	≥ 6 low scores	0.55%	0.83%		2.08%	
	≥ 7 low scores	0.20%	0%		0%	
	≥ 8 low scores	0.06%				
	≥ 9 low scores	0.02%				
	≥ 10 low scores	0.00%				
	11 low scores	0.00%				
Attentional functions (10 tests)	0 low scores	75.09%	72.50%	81.63%	61.70%	75%
	≥ 1 low score	24.92%	27.50%	18.37%	38.30%	25%
	≥ 2 low score	9.21%	14.17%	12.24%	17.02%	12.50%
	≥ 3 low scores	3.75%	7.50%**	8.16%	8.51%	4.17%
	≥ 4 low scores	1.56%	3.33%	4.08%	4.26%	0%
	≥ 5 low scores	0.51%	5–10%	5–10%	5–10%	
	≥ 6 low scores	0.15%				
	≥ 7 low scores	0.04%				
	≥ 8 low scores	0.04%				
	≥ 9 low scores					
	10 low scores					
Memory (8 tests)	0 low scores	69.44%	63.64%	79.59%	54.17%	50%
	≥ 1 low score	24.92%	36.36%	20.41%	45.83%**	50%**
	≥ 2 low score	9.64%	17.36%**	10.20%	20.83%**	25%**
	≥ 3 low scores	3.60%	7.44%	2.04%	10.42%**	12.50%
	≥ 4 low scores	1.35%	3.31%**	0%	6.25%**	4.17%
	≥ 5 low scores	0.45%	3.31%**		6.25%**	4.17%
	≥ 6 low scores	0.12%	0.83%		2.08%	0%
	≥ 7 low scores	0.03%	0%		0%	
	8 low scores	0%				
Logical reasoning (2 tests)	0 low scores	91.43%	96.69%	97.96%	93.75%	100%
	≥ 1 low scores	8.57%	3.31%	2.04%	4.17%	0%
	2 low scores	1.44%	1.65%	0%	2.08%**	

Note. Mild: patients not hospitalized for SARS-CoV-2 infection; Moderate: patients hospitalized without mechanical ventilation for SARS-CoV-2 infection; and Severe: patients hospitalized in intensive care with mechanical ventilation for SARS-CoV-2 infection.
Note. ** Cumulative percentages of patients significantly above the estimated percentage for the normative population after FDR correction.

Executive Functions

The total sample differed significantly from the normative population on one or more (+ 13.96%, $p < .001$), two or more (+ 12.91%, $p < .001$), three or more (+ 5.70%, $p = .005$), and four or more (+ 5.29%, $p = .008$) abnormally low scores, but not on five or more abnormally low scores ($p > .05$; Table 4 and Fig. 1).

The mild subgroup did not differ significantly from the normative population on any of the cumulative percentages (all ps > .05).

The moderate subgroup differed significantly from the normative population on one or more (+27.82%, $p < .001$), two or more (+ 20.62%, $p < .001$), three or more (+ 10.80%, $p = .002$), and four and more (+ 9.68%, $p < .001$) abnormally low scores, but not on five or more abnormally low scores ($p > .05$).

The severe subgroup differed significantly from the normative population on one or more (+ 18.51%, $p = .018$) and two or more (+ 16.46%, $p = .008$) abnormally low scores, but the three other differences did not survive FDR correction (all ps > .05).

Attentional Functions

The total sample differed significantly from the normative population on three or more abnormally low scores (+ 3.75%, $p = .015$), but all other cumulative percentage comparisons were nonsignificant (all ps > .05; Table 4).

For the mild, moderate, and severe subgroups, no significant differences were found after FDR correction (all ps > .05).

Memory Functions

The total sample differed significantly from the normative population on two or more (+ 7.72%, $p = .003$), four or more (+ 1.96%, $p = .001$), and five or more abnormally low scores (+ 2.86%, $p = .002$). Comparisons on one or more, three or more, and six or more abnormally low scores were nonsignificant after FDR correction (all ps > .05; Table 4 and Fig. 1).

No significant differences were found between the mild subgroup and the normative population on any of the cumulative percentages (all ps > .050).

The moderate subgroup differed significantly from the normative population on one or more (+ 20.91%, $p = .001$), two or more (+ 11.19%, $p = .009$), three or more (+ 5.82%, $p < .001$), four or more (+ 4.90%, $p < .001$), and five or more (+ 5.80%, $p = .001$) abnormally low scores.

The severe subgroup differed significantly from the normative population after FDR correction on one or more (+ 25.08%, $p = .002$) and two or more (+ 15.36%, $p = .017$) abnormally low scores, but comparisons for three or more, and five or more abnormally low scores did not survive FDR correction. All other comparisons were nonsignificant (all ps > .050).

Logical Reasoning

The total sample and mild and severe subgroups did not differ significantly from the normative population in either comparison (both ps > .05). For the moderate subgroup, a significant difference was found on two or more abnormally low scores (+ 0.65%, $p = .032$), but not on one or more abnormally low scores ($p > .05$; Table 3).

Discussion

To date, it has been difficult to determine the frequency of neuropsychological deficits in post-COVID-19 condition. Although research has highlighted a high frequency of neuropsychological deficits in various cognitive functions (Almeria et al., 2020; Jaywant et al., 2021; Krishnan et al., 2022), authors have not envisaged the possibility that this frequency could be comparable with that observed in a normative population (Binder et al., 2009). The present study was designed to address this issue, by comparing the cumulative frequency of neuropsychological deficits in patients who had had COVID-19 (mild, moderate, or severe) at 222.46 ± 42.93 days post-infection with estimated percentages for a normative population, using a robust and validated neuropsychological simulation methodology. As expected, the moderate and severe subgroups had higher cumulative percentages of memory and executive function deficits than the normative population. Moreover, our analysis suggested a greater frequency of perceptual and logical reasoning deficits in moderate (but not mild or severe) patients, compared with estimates for the normative population.

For the first time in a population of patients with no clinical history prior to the infection, our analyses showed that the frequency of neuropsychological deficits 6–9 months following SARS-CoV-2 infection was higher than would be observed in a normative population, suggesting that SARS-CoV-2 infection has long-term neuropsychological consequences. For the methodological reasons set out above, it proved difficult to compare our results with those of other studies that have assessed the
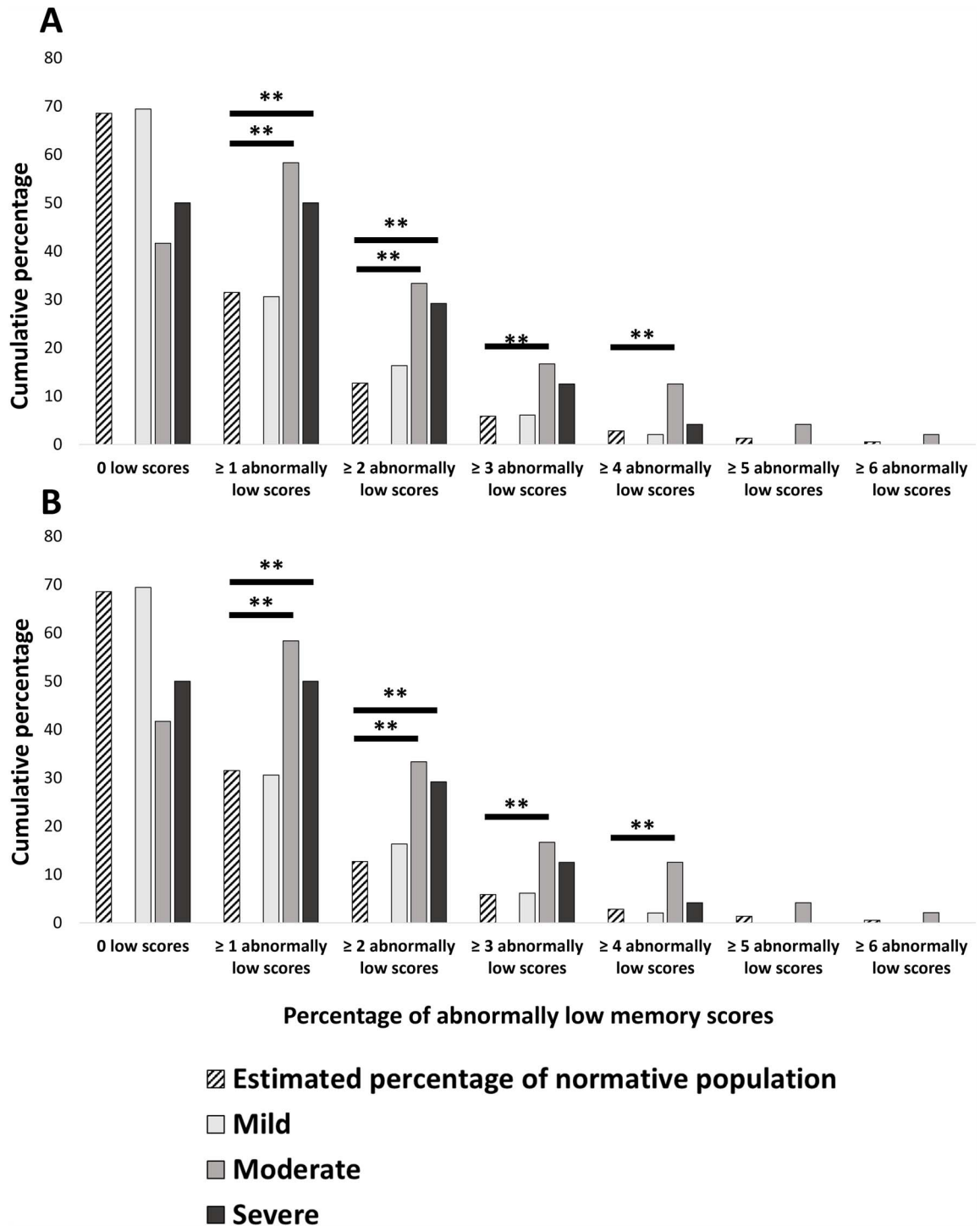
Fig. 1. Cumulative percentages for SARS-CoV-2 groups (total sample; mild, moderate, and severe subgroups) and estimated percentages for normative population, as well as FDR-corrected results of probability distribution comparisons. (A) Cumulative percentages for memory. (B) Cumulative percentages for executive functions. *Note.* Mild: patients not hospitalized for SARS-CoV-2 infection; Moderate: patients hospitalized without mechanical ventilation for SARS-CoV-2 infection; and Severe: patients hospitalized in intensive care with mechanical ventilation for SARS-CoV-2 infection. ** Cumulative percentages of patients significantly above the estimated percentage for the normative population after FDR correction.

frequency of neuropsychological deficits following SARS-CoV-2 infection. We were nevertheless able to do so to some extent in order to interpret our results. In the present study, calculating cumulative percentages of patients with neuropsychological deficits freed us from the previously mentioned methodological limitation and allowed us to compare these percentages with those estimated for a normative population, as recommended in the methodological literature on the assessment of the frequency

of deficits in neuropsychology (Crawford et al., 2007). Moreover, the neuropsychological deficits were not confined to the severe subgroup (i.e., patients who required mechanical ventilation), as they were also present in the moderate subgroup, and even appeared to be of greater severity (major perceptual and logical reasoning deficits), confirming the observations of Alemanno and colleagues (2021), who highlighted significantly greater neuropsychological deficits on the MoCA and MMSE screening batteries in moderate patients (hospitalized without mechanical ventilation) than in severe patients (hospitalized with mechanical ventilation). This suggests that post-ICU effects are not the sole cause of long-term neuropsychological deficits (Sakusic et al., 2018), given the high cumulative percentages of memory, executive, perceptual, and logical reasoning deficits in the moderate subgroup. Moreover, this highlights central nervous system involvement, either directly via SARS-CoV-2 neuroinvasive mechanisms, or indirectly via an immunological and/or inflammatory response to SARS-CoV-2 (Bougakov, Podell, & Goldberg, 2021), which is supported by PET-FDG (Guedj et al., 2021), functional connectivity (Voruz et al., 2022b), and structural (Douaud et al., 2022) neuroimaging. Another hypothesis, which could explain the greater number of cumulative deficits in moderate patients, concerns the level of hypoxemia, for according to the literature, this is a predictor of cognitive deficits (Herridge et al., 2016; Mikkelsen et al., 2012). Thus, patients who were not in intensive care and intubated may have experienced more prolonged hypoxemia, which would explain these greater neuropsychological deficits. Our results do not allow us to state exactly how the central nervous system is affected in COVID-19, but do suggest that the substantial variance in long-term neuropsychological deficits in COVID-19 is attributable to the infection, and not to previous medical history, treatment, or post-ICU effects. In contrast, our analyses failed to reveal higher cumulative percentages of deficits for the mild subgroup than for the normative population, despite a trend toward significance (nonsignificant after FDR correction) for attention. The cumulative percentages for the other functions were significantly lower, in contrast to previous observations among mild patients (Krishnan et al., 2022). Nevertheless, this does not exclude the possibility that some patients may have experienced long-term effects such as attentional or concentration problems, with an impact on their daily lives. Some patients may well have had neuropsychiatric symptoms, which are described in the literature as being very frequent (for review see Mazza et al., 2020; Schou, Joca, Wegener, & Bay-Richter, 2021). These symptoms may sporadically influence neuropsychological performance and patients' daily life, without reaching large cumulative percentages.

Limitations

The present study had five limitations. First, our subgroup of mild patients may have had a moderate or severe infection, but not have gone to hospital. This may have been the effect of sociocultural status and gender, and may have resulted in long-term cognitive deficits, owing to poor disease management. Second, although our study had strict exclusion criteria, the premorbid features (e.g., respiratory, cardiac, or endocrine pathologies) that we documented in some of our patients may have induced cognitive deficits. However, few of our patients reported comorbidities, and according to the literature, most of these comorbidities are not known to induce cognitive deficits, if treated (e.g., four of our patients reported asthma). Third, it is currently impossible for performances on a neuropsychological test to be determined by a single cognitive function. All the evidence suggests that performances are cognitively multidetermined, thereby limiting their interpretation (Bilder & Reise, 2019). Nevertheless, our pooling to measure neuropsychological subfunctions was based on validated cognitive models and validated test batteries. Fourth, a control group could have been included, but because of the high rate of infection in Geneva across 2020 and 2021 and our exclusion criteria (absence of relevant clinical history), plus the necessity of antibody testing, there was a high risk of including patients who had had a mild disease, which would have biased the study. Fifth, our analysis was an exploratory one with 121 participants. This may have influenced the generalizability of the results, and there may have been an overall tendency for the level of abnormality in each separate comparison to be overestimated (Crawford & Garthwaite, 2005). We sought to overcome this difficulty by incorporating a correction for multivariate comparisons (i.e., FDR correction), which allowed us to partially correct this tendency.

Despite these limitations, our study comparing a cohort of patients with no relevant clinical history, 6–9 months after COVID-19 infection, with a normative population created using a statistical simulation methodology validated in neuropsychology, yielded evidence of major cumulative deficits in neuropsychological functions such as memory and executive functions. These were found not only in patients who had been in ICU, but also in those who had had moderate disease, suggesting a potential direct or indirect impact of the infection on cognitive functions.

Conclusion

Our results highlighted a higher frequency of neuropsychological deficits (mainly affecting memory and executive functions) among moderate and severe patients with post-COVID-19 condition than would be expected in a normative

population. They highlight the importance of considering these symptoms in the management of patients with post-COVID-19 condition.

Funding

Conflict of Interest

None declared.

References

Abeare, K., Razvi, P., Sirianni, C. D., Giromini, L., Holcomb, M., Cutler, L., et al. (2021). Introducing alternative validity cutoffs to improve the detection of non-credible symptom report on the BRIEF. *Psychological Injury and Law*, *14*(*1*), 2–16. https://doi.org/10.1007/s12207-021-09402-4.

Alemanno, F., Houdayer, E., Parma, A., Spina, A., Del Forno, A., Scatolini, A., et al. (2021). COVID-19 cognitive deficits after respiratory assistance in the subacute phase: A COVID-rehabilitation unit experience. *PLoS One*, *16*(*2*), e0246590. https://doi.org/10.1371/journal.pone.0246590.

Almeria, M., Cejudo, J., Sotoca, J., Deus, J., & Krupinski, J. (2020). Cognitive profile following COVID-19 infection: Clinical predictors leading to neuropsychological impairment. *Brain, Behavior, & Immunity-Health*, *9*, 100163. https://doi.org/10.1016/j.bbih.2020.100163.

Benzakour, L., Braillard, O., Mazzola, V., Gex, D., Nehme, M., Perone, S. A., et al. (2021). Impact of peritraumatic dissociation in hospitalized patients with COVID-19 pneumonia: A longitudinal study. *Journal of Psychiatric Research*, *140*, 53–59. https://doi.org/10.1016/j.jpsychires.2021.05.031.

Bilder, R. M., & Reise, S. P. (2019). Neuropsychological tests of the future: How do we get there from here?. *The Clinical Neuropsychologist*, *33*(*2*), 220–245.

Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, *24*(*1*), 31–46. https://doi.org/10.1093/arclin/acn001.

Bougakov, D., Podell, K., & Goldberg, E. (2021). Multiple neuroinvasive pathways in COVID-19. *Molecular Neurobiology*, *58*(*2*), 564–575. https://doi.org/10.1007/s12035-020-02152-5.

Crawford, J. R., & Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using monte carlo simulations and revised tests for dissociations. *Neuropsychology*, *19*(*3*), 318–331. https://doi.org/10.1037/0894-4105.19.3.318.

Crawford, J. R., Garthwaite, P. H., & Gault, C. B. (2007). Estimating the percentage of the population with abnormally low scores (or abnormally large score differences) on standardized neuropsychological test batteries: A generic method with applications. *Neuropsychology*, *21*(*4*), 419–430. https://doi.org/10.1037/0894-4105.21.4.419.

Daroische, R., Hemmingyth, M. S., Eilertsen, T. H., Breitve, M. H., & Chwiszczuk, L. J. (2021). Cognitive impairment after COVID-19—A review on objective test data. *Frontiers in Neurology*, *12*. https://doi.org/10.3389/fneur.2021.699582.

Douaud, G., Lee, S., Alfaro-Almagro, F., Arthofer, C., Wang, C., McCarthy, P., et al. (2022). SARS-CoV-2 is associated with changes in brain structure in UK Biobank. *Nature*, *604*, 697–707. https://doi.org/10.1038/s41586-022-04569-5.

Drozdick, L. W., Raiford, S. E., Wahlstrom, D., & Weiss, L. G. (2018). *The Wechsler adult intelligence*. Scale—Fourth Edition and the Wechsler Memory Scale—Fourth Edition. New York, NY, US: The Guilford Press.

Eustache, F., Viard, A., & Desgranges, B. (2016). The MNESIS model: Memory systems and processes, identity and future thinking. *Neuropsychologia*, *87*, 96–109. https://doi.org/10.1016/j.neuropsychologia.2016.05.006.

Frei, A., Balzer, C., Gysi, F., Leros, J., Plohmann, A., & Steiger, G. (2016). Kriterien zur Bestimmung des Schweregrades einer neuropsychologischen Störung sowie Zuordnungen zur Funktions-und Arbeitsfähigkeit. *Neuropsychologie*, *27*, 107–119. https://doi.org/10.1024/1016-264X/a000177.

Grober, E., & Buschke, H. (1987). Genuine memory deficits in dementia. *Developmental Neuropsychology*, *3*, 13–36. https://doi.org/10.1080/87565648709540361.

Guedj, E., Campion, J., Dudouet, P., Kaphan, E., Bregeon, F., Tissot-Dupont, H., et al. (2021). 18 F-FDG brain PET hypometabolism in patients with long COVID. *European Journal of Nuclear Medicine and Molecular Imaging*, *48*(*9*), 2823–2833.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications*. In *with a supplement for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. Odessa: Psychological Assessment Resources.

Herridge, M. S., Moss, M., Hough, C. L., Hopkins, R. O., Rice, T. W., Bienvenu, O. J., et al. (2016). Recovery and outcomes after the acute respiratory distress syndrome (ARDS) in patients and their family caregivers. *Intensive Care Medicine*, *42*(*5*), 725–738. https://doi.org/10.1007/s00134-016-4321-8.

Jaywant, A., Vanderlind, W. M., Alexopoulos, G. S., Fridman, C. B., Perlis, R. H., & Gunning, F. M. (2021). Frequency and profile of objective cognitive deficits in hospitalized patients recovering from COVID-19. *Neuropsychopharmacology*, *46*, 2235–2240. https://doi.org/10.1038/s41386-021-00978-8.

Kanser, R. J., Rapport, L. J., Hanks, R. A., & Patrick, S. D. (2021). Utility of WAIS-IV Digit Span indices as measures of performance validity in moderate to severe traumatic brain injury. *The Clinical Neuropsychologist*, 1–14. https://doi.org/10.1080/13854046.2021.1921277.

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(*4*), 252–258. https://doi.org/10.1207/S15324826AN0704_8.

Krishnan, K., Miller, A. K., Reiter, K., & Bonner-Jackson, A. (2022). Neurocognitive profiles in patients with persisting cognitive symptoms associated with COVID-19. *Archives of Clinical Neuropsychology.*, *37*, 729–737. https://doi.org/10.1093/arclin/acac004.

Macoir, J., Gauthier, C., Jean, C., & Potvin, O. (2016). BECLA, a new assessment battery for acquired deficits of language: Normative data from Quebec-French healthy younger and older adults. *Journal of the Neurological Sciences*, *361*, 220–228. https://doi.org/10.1016/j.jns.2016.01.004.

Mahieux-Laurent, F., Fabre, C., Galbrun, E., Dubrulle, A., Moroni, C., & Sud, G. (2009). Validation d'une batterie brève d'évaluation des praxies gestuelles pour consultation Mémoire. Évaluation chez 419 témoins, 127 patients atteints de troubles cognitifs légers et 320 patients atteints d'une démence. *Revue Neurologique*, *165*(6–7), 560–567. https://doi.org/10.1016/j.neurol.2008.11.016.

Mazza, M. G., De Lorenzo, R., Conte, C., Poletti, S., Vai, B., Bollettini, I., et al. (2020). Anxiety and depression in COVID-19 survivors: Role of inflammatory and clinical predictors. *Brain, Behavior, and Immunity*, *89*, 594–600. https://doi.org/10.1016/j.bbi.2020.07.037.

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and recognition trial professional manual*. Odessa: Psychological Assessment Resources.

Mikkelsen, M. E., Christie, J. D., Lanken, P. N., Biester, R. C., Thompson, B. T., Bellamy, S. L., et al. (2012). The adult respiratory distress syndrome cognitive outcomes study: Long-term neuropsychological function in survivors of acute lung injury. *American Journal of Respiratory and Critical Care Medicine*, *185*(12), 1307–1315. https://doi.org/10.1164/rccm.201111-2025OC.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734.

Nehme, M., Braillard, O., Alcoba, G., Aebischer Perone, S., Courvoisier, D., Chappuis, F., et al. (2021). COVID-19 symptoms: Longitudinal evolution and persistence in outpatient settings. *Annals of Internal Medicine*, *174*(5), 723–725. https://doi.org/10.7326/M20-5926.

Roussel, M., & Godefroy, O. (2008). La batterie GREFEX: Données normatives. In O. Godefroy & GREFEX (Eds.), *Fonctions exécutives et pathologies neurologiques et psychiatriques* (pp. 231–252). Marseille: Solal.

Sakusic, A., O'Horo, J. C., Dziadzko, M., Volha, D., Ali, R., Singh, T. D., et al. (2018). Potentially modifiable risk factors for long-term cognitive impairment after critical illness: A systematic review. *Mayo Clinic Proceedings*, *93*, 68–82. https://doi.org/10.1016/j.mayocp.2017.11.005.

Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: MIT Press.

Schou, T. M., Joca, S., Wegener, G., & Bay-Richter, C. (2021). Psychiatric and neuropsychiatric sequelae of COVID-19–A systematic review. *Brain, Behavior, and Immunity.*, *97*, 328–348. https://doi.org/10.1016/j.bbi.2021.07.018.

Schretlen, D. J., Testa, S. M., Winicki, J. M., Pearlson, G. D., & Gordon, B. (2008). Frequency and bases of abnormal performance by healthy adults on neuropsychological testing. *Journal of the International Neuropsychological Society*, *14*(3), 436–445. https://doi.org/10.1017/S1355617708080387.

Voruz, P., Allali, G., Benzakour, L., Nuber-Champier, A., Thomasson, M., Jacot de Alcântara, I., et al. (2022a). Long COVID neuropsychological deficits after severe, moderate, or mild infection. *Clinical and Translational Neuroscience*, *6*(2), 9. https://doi.org/10.3390/ctn6020009.

Voruz, P., Cionca, A., Jacot de Alcântara, I., Nuber-Champier, A., Allali, G., Benzakour, L., et al. (2022b). Functional connectivity underlying cognitive and psychiatric symptoms in post-COVID-19 syndrome: Is anosognosia a key determinant? *Brain Communications*, *4*. https://doi.org/10.1093/braincomms/fcac057.

Warrington, E. K., & James, M. (1991). *The visual object and space perception battery*. Bury St. Edmunds: Thames Valley Test Company.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*, *22*(498), 1.

Zimmermann, P., & Fimm, B. (2007). *Test for attentional performance (tap), version 2.1, operating manual*. Herzogenrath: PsyTest.