scientific reports

Check for updates

OPEN

Multinational characterization of neurological phenotypes in patients hospitalized with COVID-19

Trang T. Le[1,92], Alba Gutiérrez-Sacristán[2,92], Jiyeon Son[3,92], Chuan Hong[2], Andrew M. South[4], Brett K. Beaulieu-Jones[2], Ne Hooi Will Loh[5], Yuan Luo[6], Michele Morris[7], Kee Yuan Ngiam[8], Lav P. Patel[9], Malarkodi J. Samayamuthu[7], Emily Schriver[10], Amelia L. M. Tan[2], Jason Moore[1], Tianxi Cai[2], Gilbert S. Omenn[11], Paul Avillach[2], Isaac S. Kohane[2], The Consortium for Clinical Characterization of COVID-19 by EHR (4CE)[*], Shyam Visweswaran[7,93], Danielle L. Mowery[1,93] & Zongqi Xia[3,93✉]

Neurological complications worsen outcomes in COVID-19. To define the prevalence of neurological conditions among hospitalized patients with a positive SARS-CoV-2 reverse transcription polymerase chain reaction test in geographically diverse multinational populations during early pandemic, we used electronic health records (EHR) from 338 participating hospitals across 6 countries and 3 continents (January–September 2020) for a cross-sectional analysis. We assessed the frequency of International Classification of Disease code of neurological conditions by countries, healthcare systems, time before and after admission for COVID-19 and COVID-19 severity. Among 35,177 hospitalized patients with SARS-CoV-2 infection, there was an increase in the proportion with disorders of consciousness (5.8%, 95% confidence interval [CI] 3.7–7.8%, $p_{FDR} < 0.001$) and unspecified disorders of the brain (8.1%, 5.7–10.5%, $p_{FDR} < 0.001$) when compared to the pre-admission proportion. During hospitalization, the relative risk of disorders of consciousness (22%, 19–25%), cerebrovascular diseases (24%, 13–35%), nontraumatic intracranial hemorrhage (34%, 20–50%), encephalitis and/or myelitis (37%, 17–60%) and myopathy (72%, 67–77%) were higher for patients with severe COVID-19 when compared to those who never experienced severe COVID-19. Leveraging a multinational network to capture standardized EHR data, we highlighted the increased prevalence of central and peripheral neurological phenotypes in patients hospitalized with COVID-19, particularly among those with severe disease.

The World Health Organization declared coronavirus disease 2019 (COVID-19) due to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection as a global pandemic on March 11, 2020[1]. Growing evidence points to the multi-organ involvement of COVID-19, particularly the nervous system, which increases morbidity and mortality[2–4]. Given the health consequences of neurological complications, recognizing the

[1]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. [2]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [3]Department of Neurology, University of Pittsburgh, Biomedical Science Tower 3, Suite 7014, 3501 5th Avenue, Pittsburgh, PA 15260, USA. [4]Department of Pediatrics, Wake Forest School of Medicine, Winston Salem, NC, USA. [5]Department of Critical Care, National University Health Systems, Singapore, Singapore. [6]Department of Preventive Medicine, Northwestern University, Chicago, IL, USA. [7]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA. [8]Department of Surgery, National University Health Systems, Singapore, Singapore. [9]Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS, USA. [10]Data Analytics Center, University of Pennsylvania Health System, Philadelphia, PA, USA. [11]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. [92]These authors contributed equally: Trang T. Le, Alba Gutiérrez-Sacristán and Jiyeon Son. [93]These authors jointly supervised this work: Shyam Visweswaran, Danielle L. Mowery and Zongqi Xia. [*]A list of authors and their affiliations appears at the end of the paper. ✉email: zxia1@post.harvard.edu
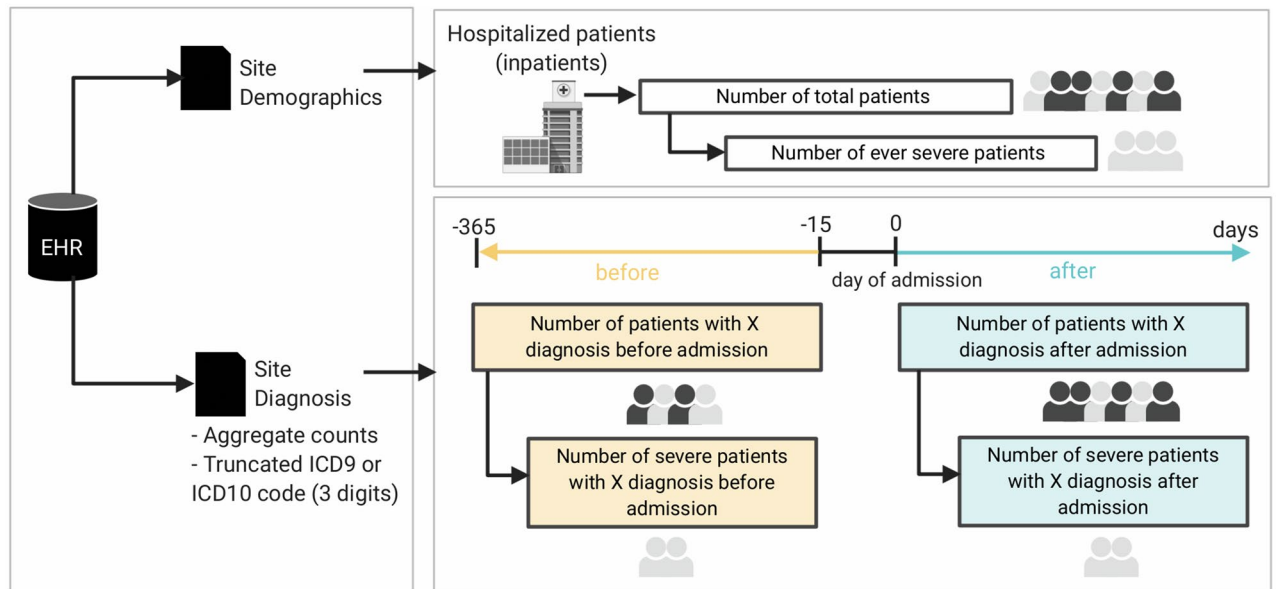
Figure 1. Schematic diagram of the cohort and data generation workflow for each healthcare system. The figure was created with Biorender (Biorender.com).

neurological phenotypes associated with COVID-19 would inform prevention, diagnosis and treatment that could potentially mitigate disability and death.

Early reports highlighted central and peripheral neurological phenotypes in adults with COVID-19, including cerebrovascular disease[5], meningoencephalitis and encephalomyelitis[6,7], encephalopathy[5], cranial neuropathies[8], Guillain-Barré syndrome[9–11], plexopathy[12], anosmia and ageusia[13,14] and cognitive and neuropsychiatric issues[15]. Children with COVID-19 have similar presentations, including ischemic stroke, encephalopathy, headache, and muscle weakness[16–18]. Prior prevalence studies of neurological conditions in COVID-19 largely examined data from single countries (*e.g.,* China[19], the United Kingdom[20,21] and Italy[22]) or single healthcare systems[4,12,23–25]. Few *large-scale* studies have used a *standardized* data capture approach to examine the multinational prevalence of the spectrum of neurologic conditions in COVID-19 patients and with *careful local* data quality control, particularly those who experienced severe respiratory and/or critical illness status[26,27].

Electronic health records (EHRs) data can facilitate clinical discovery efforts. Our team created the International Consortium for Clinical Characterization of COVID-19 by EHR (4CE; http://www.covidclinical.net) to standardize and aggregate multinational EHR data (from 34 healthcare systems and 338 affiliated hospitals across six countries at the time of data freeze for this study) to address critical clinical and epidemiological questions relevant to COVID-19[28–31]. Central to the 4CE effort is the ability of local clinician experts and data scientists at each contributing healthcare system to ensure the quality of common EHR data elements. Leveraging the highly scalable, federated, multinational networks of the 4CE consortium, we computed the prevalence of a wide range of central and peripheral neurological conditions in hospitalized patients with reverse transcription polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection by healthcare system and by country. We also compared the differences in the prevalence of neurological conditions between patients with and those without severe COVID-19 based on the internationally validated 4CE severity criteria[30].

Results

Following a consortium-wide standardized procedure (Fig. 1), we collected the EHR data from 35,177 hospitalized patients with PCR-confirmed SARS-CoV-2 infection from 338 hospitals affiliated with 34 healthcare systems in six countries (eTable 1). Aggregate demographic data were available for 34,647 patients (98.5%). The cohort had a greater proportion of men (20,814, 60.1%) than women (13,546, 39.1%), while 287 (0.8%) patients had unknown gender (Fig. 2A). The study captured a broad range of proportions of severe COVID-19 cases (based on the internationally validated 4CE COVID-19 severity criteria[30]) across this multinational network of healthcare systems. There was no clear relationship between COVID-19 severity and median age (Fig. 2B). Most healthcare systems in Europe did not report race. Among the US healthcare systems, there was a disproportionately high proportion of self-reported Black individuals (Fig. 2C). The study population included a high proportion of individuals above age 50 years and a low proportion of children (age < 18 years) (Fig. 2D).

We first assessed the prevalence of a wide spectrum of neurological conditions during the first hospital admission for COVID-19, importantly, using the estimates in the 1-year pre-admission period at each healthcare system as the comparator. Towards this end, we used EHR data from the early phase of the pandemic (January to September 2020) and queried all potential neurological conditions based on a comprehensive literature review at the start of the analysis. Most of the contributing healthcare systems (77%) reported an increase in the proportion of hospitalized COVID-19 patients with disorders of consciousness (ICD-10 R41: "Other symptoms and signs involving cognitive functions and awareness") with a mean increase of 5.8% (95% CI 3.7–7.8%, $p_{FDR} < 0.001$)
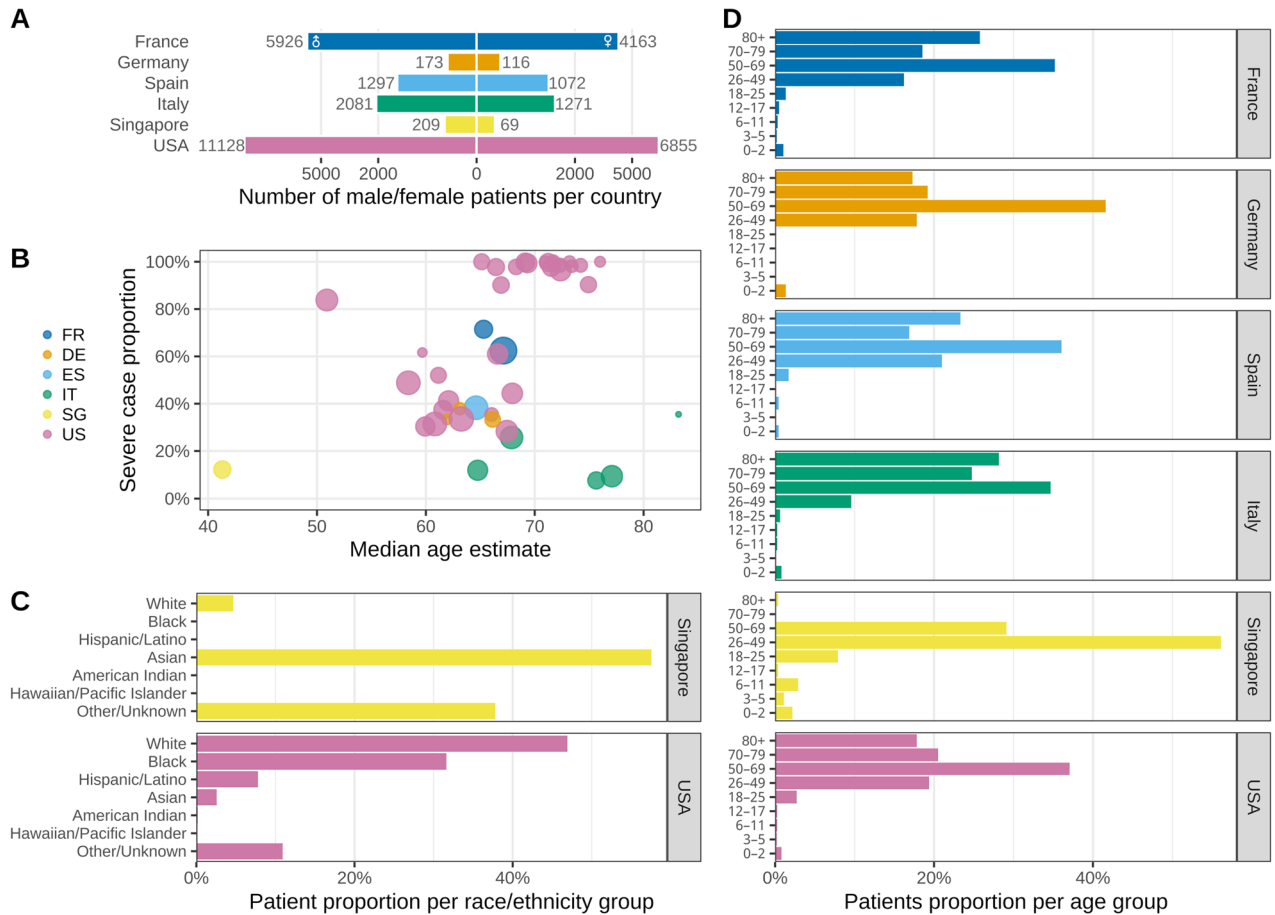
Figure 2. Characteristics of the study population across healthcare systems and countries. (**A**) Total number of male (left) and female (right) patients grouped by country shown in square-root scale. (**B**) Proportion of ever-severe cases by median age estimate at each healthcare system, grouped by country. Node size corresponds to the total number of patients per system. (**C**) Distribution of self-identified race among patients at healthcare systems in Singapore and the United States. The Other/Unknown category includes patients who did not identify with any of the predefined race categories and/or whose data were not reported. Most European healthcare systems did not report race. (**D**) Average proportion of patients in each age group within each country. FR, France; DE, Germany; ES, Spain; IT, Italy; SG, Singapore; US(A), United States of America.

after admission (Fig. 3, eFig. 1). Similarly, 84% of healthcare systems reported an increase in the proportion of patients with "Other disorders of the brain" (ICD-10 G93, including "encephalopathy", "cerebral edema", "brain death") with a mean increase of 8.1% (5.7–10.5%, $p_{FDR} < 0.001$) after admission (see online interactive data repository: https://covidclinical.github.io/Phase1.1NeuroRCode/01-analysis-icd10.html#prevalence-change-table). The proportion of patients with "epilepsy and recurrent seizures" (ICD-10 G40), "encephalitis, myelitis, and encephalomyelitis" (ICD-10 G04) and "other and unspecified myopathies" (ICD-10 G72) increased after admission, but these findings were not significant after adjusting for multiple testing. Likewise, none of the other neurological conditions showed a statistically significant difference in prevalence after admission.

To assess the association with COVID-19 severity, we next used Fisher's exact test to examine the enrichment or depletion of each neurological condition among patients hospitalized for COVID-19 who ever experienced severe disease based on the published 4CE COVID-19 severity criteria[30], using those who never experienced severe disease as the comparator (Fig. 4). A positive $\log_2$ value of enrichment (LOE) value denoted a higher proportion of severe cases for a given neurological condition than the never-severe cases, while a negative LOE value indicated the opposite. For instance, a LOE value of 0.283 for ICD code R41 meant that the observed number of severe cases with R41 (disorder of consciousness) was $2^{0.283}$ or ~ 1.22 times higher than the expected number of severe cases for R41, which was equivalent to a 22% increase in relative risk (*i.e.,* relative risk difference RRD = 22%). Table 1 listed the neurological phenotypes that exhibited *statistically significant* associations with severe COVID-19 status ($p_{FDR} < 0.05$). The interactive data table (https://covidclinical.github.io/Phase1.1NeuroRCode/01-analysis-icd10.html#enrichment_tab) and the results directory of the project online data repository (https://github.com/covidclinical/Phase1.1NeuroRCode/tree/master/results) listed the LOE, 95% confidence intervals and *p* values for *all* neurological conditions examined.

In the period after hospital admission for COVID-19, using the patients who never experienced severe COVID-19 as reference, a significantly higher proportion of patients with severe disease had "other symptoms and signs involving cognitive functions and awareness" (ICD-10 R41: RRD$_{after}$ = 22%), "other cerebrovascular
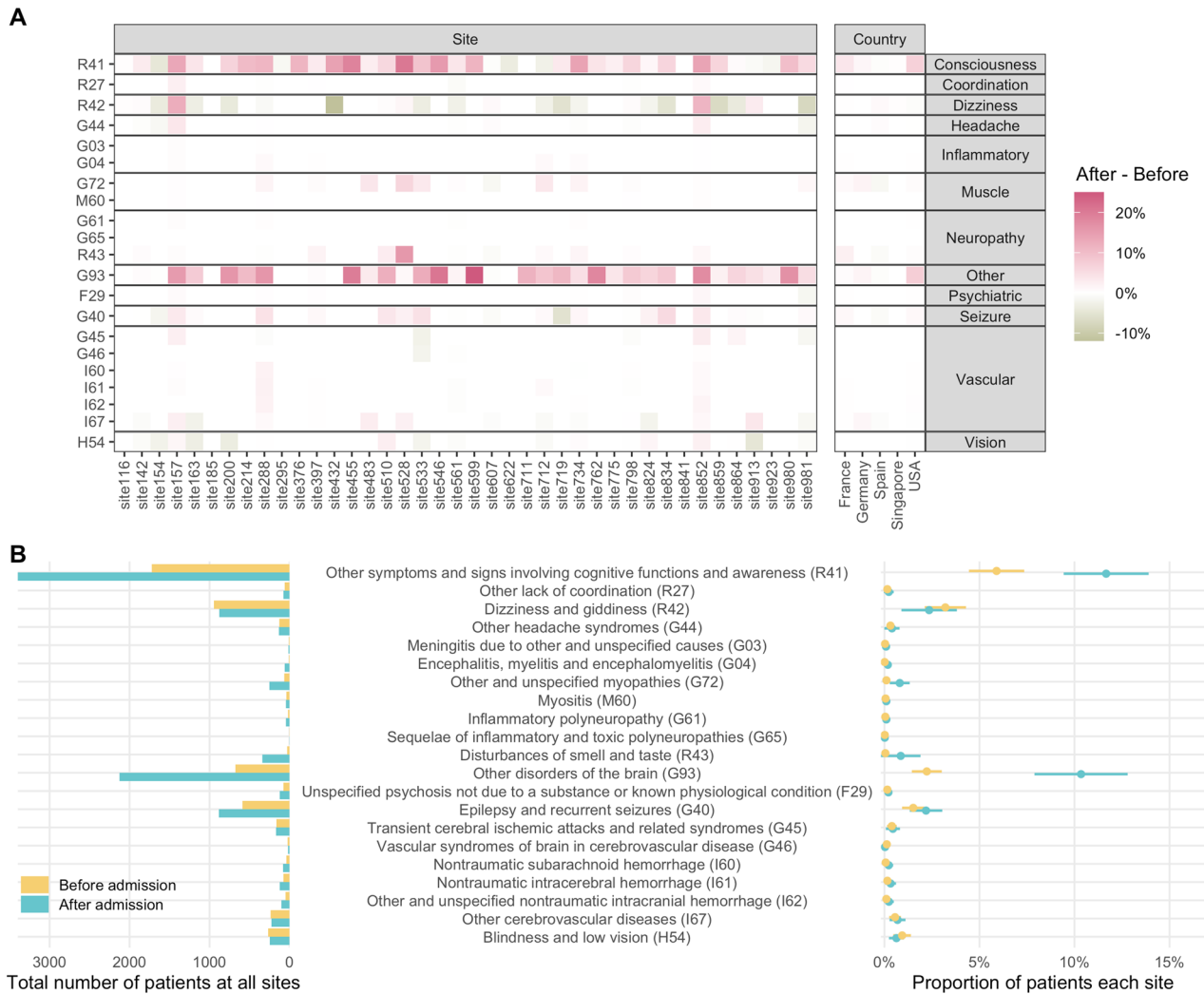
Figure 3. Prevalence of neurological phenotypes among all patients. (**A**) Difference in prevalence of each neurological ICD-10 code by healthcare system and country, calculated as after admission—before admission date (eEq. 2). Pink color on the heat map indicates increased prevalence, while green color indicates decreased prevalence. Please see eFig. 1 for the absolute values of prevalence. (**B**) Total counts of patients with a given neurological ICD-10 code (left) and the mean proportion of patients (right) before and after admission date across all healthcare systems. The mean proportion estimates are shown as circles and the 95% confidence intervals are shown as bars.

disease" (ICD-10 I67: $RRD_{after} = 24\%$), "nontraumatic subarachnoid hemorrhage" (ICD-10 I60: $RRD_{after} = 28\%$), "other and unspecified nontraumatic intracranial hemorrhage" (ICD-10 I62: $RRD_{after} = 34\%$), "nontraumatic intracerebral hemorrhage" (ICD-10 I61: $RRD_{after} = 36\%$), "other disorders of the brain" (ICD-10 G93, including "encephalopathy": $RRD_{after} = 36\%$), "encephalitis, myelitis and encephalomyelitis" (ICD-10 G04: $RRD_{after} = 37\%$), and "other and unspecified myopathies" (ICD-10 G72, including "inflammatory and immune myopathies" and "critical illness myopathies": $RRD_{after} = 72\%$) (Table 1, Fig. 4). In contrast, a significantly lower proportion of patients with severe disease had "blindness and low vision" (ICD-10 H54: $RRD_{after} = -23\%$), "dizziness and giddiness" (ICD-10 R42: $RRD_{after} = -28\%$), "other headache syndromes" (ICD-10 G44: $RRD_{after} = -47\%$), "transient cerebral ischemic attacks and related syndromes" (ICD-10 G45: $RRD_{after} = -45\%$), and "unspecified psychosis not due to a substance or known physiological condition" (ICD-10 F29: $RRD_{after} = -44\%$) during hospitalization. Subgroup analysis comparing the US and non-US healthcare systems yielded consistent results (eFig. 2).

As 12 healthcare systems in Italy and the US contributed EHR data entirely or partially comprising ICD-9 codes, we performed separate subgroup analyses using only the ICD-9 data (eTable 2), given that one-to-one mapping of some ICD-9 codes to ICD-10 codes was not feasible. To summarize, the prevalence of "disorders of consciousness" and "other neurological conditions" increased after COVID-19 admission date when compared to the 1-year pre-admission period, as in the ICD-10 data analysis (eFigs. 3, 4). There was no statistically significant difference when examining the change in prevalence of other neurological ICD-9 codes after admission date. In severity analysis, there were similarities with the ICD-10 data (*e.g.,* "disorders of consciousness") but also differences involving seizure and cerebrovascular events that would require caution in interpretation due to differences in sample size (see Supplementary Material, eFig. 5, eTable 3).
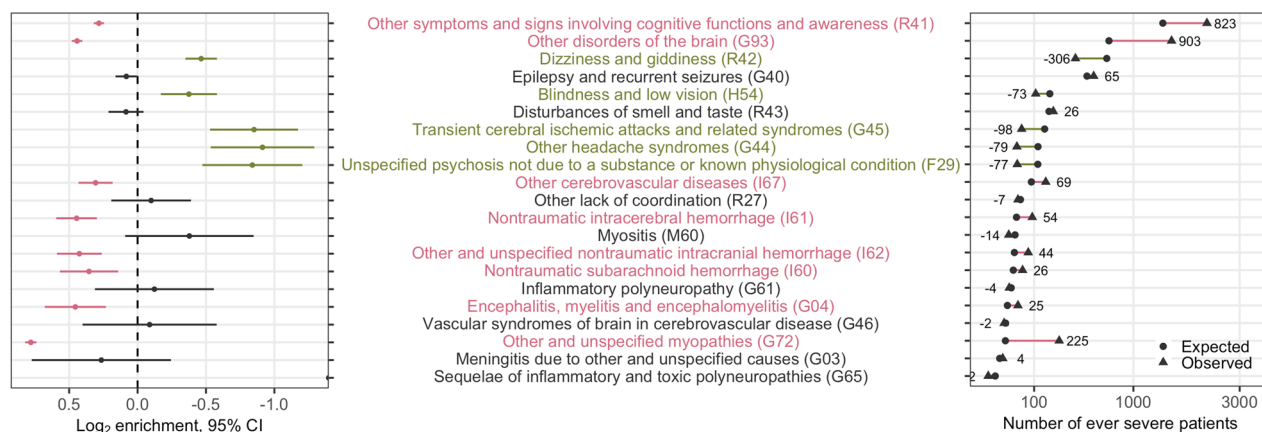
Figure 4. Analysis of enrichment or depletion of neurological conditions after admission in patients with severe disease. For each neurological ICD-10 code, we show the $\log_2$ enrichment (LOE) and its 95% confidence interval (left), and the absolute difference between the observed (filled triangle) and expected (·) number of patients experiencing severe COVID-19 in square-root scale (right). A purple positive LOE value for an ICD-10 code indicates a statistically significantly higher proportion of severe cases having a given neurological ICD-10 code when compared to the never-severe cases. Conversely, a green negative LOE value indicates a statistically significantly lower proportion of severe cases having a given neurological ICD-10 code when compared to the never-severe cases. Neurological ICD-10 codes are ordered by the expected number of severe cases after admission.

Neurological condition (ICD-10 code)	LOE[a]	RRD[b] (%)	RRD 95% CI (%)	p_{FDR}
Blindness and low vision (H54)	− 0.38	− 23	(− 33, − 11)	2.0×10^{-4}
Dizziness and giddiness (R42)	− 0.47	− 28	(− 33, -22)	6.3×10^{-19}
Encephalitis, myelitis and encephalomyelitis (G04)	0.45	37	(17, 60)	0.0081
Nontraumatic intracerebral hemorrhage (I61)	0.45	36	(23, 51)	3.6×10^{-5}
Nontraumatic subarachnoid hemorrhage (I60)	0.35	28	(10, 48)	0.019
Other and unspecified myopathies (G72)	0.78	72	(67, 77)	8.8×10^{-45}
Other and unspecified nontraumatic intracranial hemorrhage (I62)	0.43	34	(20, 50)	3.0×10^{-4}
Other cerebrovascular diseases (I67)	0.31	24	(13, 35)	2.0×10^{-4}
Other disorders of the brain (G93)	0.44	36	(32, 40)	5.6×10^{-73}
Other headache syndromes (G44)	− 0.91	− 47	(− 59, − 31)	9.4×10^{-9}
Other symptoms and signs involving cognitive functions and awareness (R41)	0.28	22	(19, 25)	2.1×10^{-39}
Transient cerebral ischemic attacks and related syndromes (G45)	− 0.85	− 45	(− 56, − 31)	5.0×10^{-10}
Unspecified psychosis not due to a substance or known physiological condition (F29)	− 0.84	− 44	(− 57, − 28)	7.7×10^{-8}

Table 1. Statistically significant associations of neurological conditions and severe disease status after admission ($p_{FDR} < 0.05$). [a]The interactive data table (https://covidclinical.github.io/Phase1.1NeuroRCode/01-analysis-icd10.html#enrichment_tab) and the results directory of the project online data repository (https://github.com/covidclinical/Phase1.1NeuroRCode/tree/master/results) show the $\log_2$ value of enrichment (LOE), 95% confidence intervals, and p values for *all* neurological ICD-codes adjusted for multiple hypothesis testing. [b]RRD: Relative risk difference = Observed relative risk − 1.

Discussion

Following a standardized approach for aggregating EHR-derived clinical facts from a federated network of multinational healthcare systems while preserving patient privacy, we report the change in prevalence of a wide spectrum of central and peripheral neurological conditions among a large cohort of patients hospitalized with PCR-confirmed SARS-CoV-2 infection across geographically diverse healthcare systems from six countries, using the 1-year period before COVID-19 hospitalization as reference. We further report the difference in the prevalence of neurological conditions among hospitalized patients with severe COVID-19 when compared to the non-severe patients. Given the challenges with evaluating patients and obtaining data during the pandemic, EHR-based studies complement the more labor-intensive physician-reported registry studies.

Disorders of consciousness and other disorders of the brain are the most prevalent neurological phenotypes among all patients hospitalized with COVID-19 during the early pandemic. While the specific underlying causes could be broad (*e.g.,* metabolic disturbance, hypoxia, medication effect, seizures, stroke), these findings are

consistent with prior reports and associated with worse health outcomes[19,20,24,25]. Given that hospitalized COVID-19 patients can manifest a wide range of symptoms and signs, certain neurological conditions do not cleanly fall within parent diagnostic codes with clearly specific descriptions (*e.g.,* ICD-10 G04: encephalitis, myelitis and encephalomyelitis; ICD-10 G61: inflammatory polyneuropathy). In the case of "other disorders of brain" (including "unspecified disorders of the brain"), the parent diagnostic code of ICD-10 G93 covers a wide range of conditions, some of which are potentially relevant to hospitalized COVID-19 patients though they do not belong to a single unifying category (*e.g.,* G93.1: anoxic brain damage; G93.2: benign intracranial hypertension; G93.3: post-viral fatigue syndrome; G93.4: other and unspecified encephalopathy; G93.5: compression of brain; G93.6: cerebral edema; G93.82: brain death). Other diagnosis codes under G93 (*e.g.,* G93.9: unspecified disorders of the brain) likely capture symptoms and signs that may *not* be easily categorized, commonly reported (*i.e.,* symptoms and signs that are *not* headache, visual changes, dizziness, confusion, weakness, or sensory changes), and/or revealing any objective pathology based on common clinical tools such as physical exam, laboratory, neuroimaging or electrodiagnostic tests. Here are two potential clinical scenarios that could lead to the diagnosis code of "unspecified disorder of a brain": (1) non-specific findings on the brain magnetic resonance imaging that were not consistent with ischemic stroke, intracranial hemorrhage, meningitis or encephalitis, neoplasm, or demyelination; (2) subjective sensory complaint without objective findings on exam or diagnostic tests that was not coded with specific diagnosis. In contrast, the diagnosis code ICD-10 R41 ("disorders of consciousness") specifically refers to symptoms and signs involving awareness and cognitive function, which include disorientation, amnesia, neglect, age-related cognitive decline, altered mental status, specific cognitive deficits involving attention, concentration, communication, visuospatial and psychomotor domains. This code will cover common clinical scenarios in hospitalized patients such as delirium (fluctuating arousal state) and decreased responsiveness.

On the other hand, we find no statistically significant increase in the prevalence of other previously reported neurological conditions such as dizziness, headache, seizure, vascular, and vision disorders after the first admission to the hospital for COVID-19. The lower prevalence of early COVID-19 symptoms such as alterations in smell and taste[13] in this study is likely attributable to the incomplete documentation of these symptoms in the EHR for the hospitalized patient population, particularly those with severe COVID-19. Crucially, these discrepancies may also be due to methodological differences, as our analysis accounted for baseline pre-admission prevalence by reporting the change in prevalence after admission.

Our other major finding indicates that a significantly higher proportion of hospitalized patients with severe COVID-19 (based on a computational phenotyping algorithm of COVID-19 severity that our group previously published[30]) had disorders of consciousness and other disorders of the brain, encephalitis and/or myelitis, cerebrovascular events and myopathy when compared to patients who never had severe disease. Beyond corroborating prior reports[19,23–25], our findings highlighted similar patterns across geographically diverse multinational healthcare systems using a standardized approach. First among these findings, disorders of consciousness and other disorders of the brain include altered mental status, disorientation cognitive deficits, and encephalopathy. Hypoxemia from respiratory failure, metabolic disturbance, sedation for advanced respiratory support, acute delirium, and other more specific neurological involvement in the setting of severe COVID-19 could all be contributory[32].

Second, encephalitis and myelitis have variable manifestations[12] and were likely under-reported due to difficulties of performing diagnostic studies (*e.g.,* magnetic resonance imaging, lumbar puncture) especially during the early phase of pandemic. Neuropathological examinations have not uncovered evidence of direct viral infection of the central nervous system (CNS)[33], though more studies are needed to confirm whether the mechanisms underlying encephalitis (with or without myelitis) are direct CNS invasion by SARS-CoV-2, acute systemic inflammation with secondary CNS involvement, and/or post-infectious immune-mediated effect on the CNS[2,34,35].

Third, our finding of cerebrovascular diseases associated with severe COVID-19 include both ischemic strokes and intracranial hemorrhages (nontraumatic intracerebral and nontraumatic subarachnoid), highlighting the difficult balance when managing severe COVID-19 with respect to antiplatelet and anticoagulation therapy. Strokes that occurred in the setting of COVID-19 were associated with high mortality and morbidity[36]. COVID-19 might increase the risk of ischemic stroke through mechanisms such as activation of innate immune system, cardioembolic events, hypoxia-induced ischemia secondary to severe pulmonary disease, coagulation activation, thrombotic angiopathy and endothelial damage[2]. Proposed mechanisms underlying intracranial hemorrhage in COVID-19 include coagulation abnormalities, endothelial dysfunction, dysregulation of the renin-angiotensin system, and disruption of cerebral blood flow autoregulation[2,37]. Patients with severe COVID-19 may have additional risk factors such as hypertension or cardiovascular disease that could further drive cerebrovascular diseases[38].

Finally, myopathy is common among severe COVID-19 patients. The cause is likely multifactorial: prolonged or severe critical illness, complications due to multi-organ involvement, or medication-induced myotoxicity (*e.g.,* Hydroxychloroquine, steroids)[2,39]. The current data set does not permit subgroup analysis (*e.g.,* ventilated versus non-ventilated patients). Further, most clinical studies to date, including our own, cannot establish whether the neurological phenotypes associated with severe COVID-19 are the direct consequence of SARS-CoV-9 neurotoxicity or due to secondary causes.

Interestingly, some neurological phenotypes are less prevalent in patients with severe COVID-19, including psychosis, dizziness, vision impairment, transient ischemic attack, and headache. The likely explanation is that critically ill patients with or without respiratory failure would either not have an objective evaluation or proper documentation and coding for these conditions given that patients with severe disease are likely to be sedated and/or having altered mental status.

Several analytical elements strengthen the study. Chief of among them, the federated approach of adopting common EHR data elements and standardized processes for representing clinical events with *local* quality control

differentiates this study from other EHR-based efforts, complements physician/neurologist-reported registry efforts[12,24,25,40] and is well suited for multinational and multi-institutional clinical discovery. Critically, the local clinician experts and data scientists at each 4CE contributing healthcare system ensure the control of EHR data quality according to the consortium standard and improve the study rigor. The overall concordance of the main study findings with other registry-based studies is reassuring. Second, our approach of aggregating EHR data, specifically the ICD codes at categorical level, reduced the potential concern for variations in coding practice across diverse healthcare systems that include general medical, neurology, critical care, and rehabilitation settings. Third, the 4CE contributing healthcare systems shared aggregate clinical data following a pre-defined analysis plan while adhering to multi-national patient privacy laws such as the United States Health Insurance Portability and Accountability Act (HIPAA) and the European Union General Data Protection Regulation (GDPR). Fourth, we rapidly implemented the analysis plan at scale by leveraging existing informatics infrastructures and frameworks at each 4CE contributing healthcare system. This federated approach of using EHR data for clinical discovery presents a complementary and alternative approach to the more labor-intensive registry-based approach. Finally, our prevalence study used the 1-year pre-admission period for comparison, which generated more realistic prevalence estimation of neurological conditions among hospitalized patients with COVID-19 than approaches that do not account for pre-admission prevalence. Given the protracted and changing nature of the COVID-19 pandemic, we intentionally included only data during the first hospitalization for COVID-19, specifically during the early phase of the pandemic, to appropriately account for pre-admission prevalence.

Our study has limitations as the result of trade-offs to standardize common data collection from multinational healthcare systems while strictly preserving patient privacy and adhering to privacy laws governing all contributing healthcare systems. First, this study relied on ICD codes that may not capture fully or accurately the disease phenotypes, particularly for conditions better documented in clinical notes. To standardize the collection of ICD codes across diverse contributing healthcare systems and to mitigate coding discrepancies, we used ICD codes at the categorical level (*e.g.*, the first 3 alphanumeric characters before the decimal point for ICD-10). As such, further characterization of certain conditions such as "other disorders of the brain" was not feasible at this stage, though we are working towards a standardized approach of capturing full ICD codes for the next stage of analyses. Second, because we aggregated data across healthcare systems, we were unable to consolidate all related ICD codes (*e.g.*, organizing into PheCode[41]) at the individual patient level. Similarly, the 4CE consortium is preparing a standardized approach to enable aggregate patient-level analyses from each contributing healthcare system. Third, we might not have captured all pre-admission EHR data if patients did not receive their entire care in the same hospital system as the COVID-19 admission. This is a limitation common to all research using EHR data from countries without universal health systems such as the USA. Reassuringly, subgroup analysis showed consistent results between US and non-US contributing healthcare systems. Finally, contributing healthcare systems with small patient counts used obfuscation to reduce the risk of re-identifying the health systems, but the effect of obfuscation is negligible because few healthcare systems had neurological conditions of interest in cases fewer than the obfuscation level. Despite the limitations of deploying this rapid, scalable, patient privacy-preserving research strategy, our key findings were consistent with prior reports from well-characterized but often smaller, single-country and/or single-center cohort studies.

In conclusion, this multinational prevalence study highlighted a range of central and peripheral neurological phenotypes in hospitalized patients with PCR-confirmed SARS-CoV-2 infection, particularly among patients with severe disease. Our multinational and multi-institutional EHR-based efforts using a standardized procedure and common data elements with careful local data quality control complement registry-based research design. In future studies, we will conduct individual-level analysis using additional EHR data such as complete ICD codes, identify risk factors for worse health outcomes (*e.g.*, hospitalization duration, death, re-admission) and examine long-term sequelae in COVID-19 patients with neurological phenotypes.

Methods

Patients and data.   4CE contributing healthcare systems began in March 2020 to collect EHR data from hospitalized patients with positive SARS-CoV-2 RT-PCR tests. The analyzed data captured the early phase of the pandemic, spanning from January 2020 through early September 2020. We defined COVID-19-related hospitalization as the first hospital admission that occurred between 7 days before and up to 14 days after the first positive SARS-CoV-2 PCR test. The first admission date within this − 7 to + 14 day window is the index admission date.

According to the 4CE consortium agreement, we de-identified contributing healthcare systems to protect their confidentiality. The institutional review board of each participating health system (eTable 1) approved the sharing of anonymous, aggregate data in compliance with multi-national patient privacy laws exempting the requirement for individual patient consent as there was no direct patient recruitment or contact. Some healthcare systems applied a small level of obfuscation (*i.e.*, masking of low counts, eTable 1) to preserve system-specific privacy and to reduce the risk of patient re-identification, though it had no significant impact on the total patient counts.

Using the 4CE standard of common EHR data elements[28–31], we collected demographics (age, gender, self-identified race/ethnicity) and the International Classification of Disease (ICD) codes (versions 9 or 10) pertaining to neurological conditions (Fig. 1) as well as COVID-19 severity, based on the internationally validated 4CE COVID-19 severity criteria[30]. Among the contributing healthcare systems, only Italian healthcare systems provided exclusively ICD-9 codes while the rest of the healthcare systems provided predominantly ICD-10 codes. As such, we used ICD-10 data for the main analyses and ICD-9 data for supplementary analyses. To standardize EHR data elements, we used the first three alphanumeric characters of a given ICD code, which designates the category of the disease or injury (*e.g.*, ICD-10 G44 denotes the category of "Other headache syndromes").

Severe illness category	Clinical events
Diagnoses	Acute respiratory distress syndrome, ventilator-associated pneumonia
Procedures	Insertion of endotracheal tube; invasive mechanical ventilation
Laboratory results	$PaCO_2$, PaO_2
Medications	General anesthetics; benzodiazepine derivatives; muscle relaxants; other hypnotics and sedatives; adrenergic and dopaminergic agents; other cardiac stimulants; other respiratory system products; phosphodiesterase inhibitors; platelet aggregation inhibitors excluding heparin; vasopressin and analogues

Table 2. The 4CE criteria of severe COVID-19. Comprising the occurrence of diagnoses, procedures, laboratory results and medications, this computational phenotyping algorithm of severity has been internationally validated (through manual chart review by local clinician experts at participating healthcare systems) to be a clinically reasonable proxy for hospitalized patients who experienced severe status of COVID-19. This approach is applicable when the *aggregate* electronic health records data from each contributing healthcare systems are available but not the patient-level data. Please see Methods and further detail in a separate 4CE publication[30].

For all patients hospitalized with COVID-19, we collected ICD codes at two time periods, *before* and *after* the date of admission of the first hospitalization for COVID-19. The period before COVID-19 hospitalization ranged from − 365 days to − 15 days preceding the admission date. Inclusion of EHR data up to 1 year before admission is a pragmatic decision to balance available data from all healthcare systems and minimize past medical conditions that might not be relevant. The period after admission date ranged from the date of admission to the end of the hospitalization. According to the pre-planned consortium-wide strategy, we excluded all codes in the 2 weeks preceding the index hospital admission date to ensure that diagnoses before admission were independent of COVID-19. It could take a few days from SARS-CoV-2 infection to symptom onset and additional days before a positive PCR test and/or hospital admission. Similarly, we analyzed ICD codes before and after the admission date according to whether patients ever met the 4CE criteria for severe COVID-19[30] (Fig. 1).

Exposures. We first examined all hospitalized patients with PCR-confirmed positive SARS-CoV-2. We then examined hospitalized patients with COVID-19 who met the 4CE criteria for severe COVID-19, including advanced respiratory care management at any point during their hospitalization. Including diagnoses, procedures, laboratory results and medications (Table 2). Because we used aggregate EHR data in this study, patient-level indicators of severity (*e.g.,* patient-level laboratory value or medication) were unavailable. In response, we applied a computational algorithm of COVID-19 severity that the 4CE consortium previously developed and internationally validated (through chart review by *local* clinician experts at participating healthcare systems) to be a clinically reasonable proxy for hospitalized patients who experienced severe COVID-19 status[30].

Neurological outcomes. We queried 21 ICD-10 codes in 12 categories pertaining to neurological phenotypes following a comprehensive literature search in August 2020. The neurological disease categories included consciousness, coordination, dizziness, headache, inflammatory, muscle, neuropathy, psychiatric, seizure, vascular, vision and other unspecified neurological conditions (Table 3).

Statistical analyses. We first compared the prevalence of each neurological ICD code and disease category among all hospitalized patients with COVID-19. For each ICD code, we reported the total count and the proportion of patients hospitalized with COVID-19 at each healthcare system (and each country), both before and after admission date (Supplementary Material, eEq. 1). We used the proportion data before admission as reference control. We calculated the difference in the proportion of cases with a given ICD code before and after admission date (eEq. 2) and used paired two-sided t-tests to examine whether there was a statistically significant difference in the proportion after admission date when compared to that before admission date.

We next compared the prevalence of each neurological ICD code and disease category after admission date between patients who ever met the criteria of severe COVID-19 and those who did not, by healthcare system and by country. For each ICD code, we computed the expected number of severe cases (eEq. 3) and compared with the observed number of severe cases for the given neurological code. To examine the difference in proportion of severe cases for the neurological ICD codes, we calculated the enrichment of each neurological ICD code by dividing the observed number of severe cases by the expected number of severe cases and reported a value of $\log_2$ enrichment (LOE) and its 95% confidence interval (CI) (eEq. 4). We estimated the LOE 95% CI using the Delta method[42]. We chose LOE as a statistic measure for the difference between the proportion of severe cases and never-severe cases for a given neurological condition because it allows symmetric visualization of enrichment or depletion (equivalent confidence intervals for enrichment and depletion. Finally, we computed the *p*-values using Fisher's exact test[43] and corrected for multiple hypothesis testing with Benjamini–Hochberg's false discovery rate (FDR) procedure[44]. A result was statistically significant if $p_{FDR} < 0.05$. All analyses were performed in the R environment.

Disease category	ICD-10 code[a]	ICD-10 code description
Consciousness	R41	Other symptoms and signs involving cognitive functions and awareness
Coordination	R27	Other lack of coordination
Dizziness	R42	Dizziness and giddiness
Headache	G44	Other headache syndromes
Inflammatory	G03	Meningitis due to other and unspecified causes
Inflammatory	G04	Encephalitis, myelitis and encephalomyelitis
Muscle	G72	Other and unspecified myopathies
Muscle	M60	Myositis
Neuropathy	G61	Inflammatory polyneuropathy
Neuropathy	G65	Sequelae of inflammatory and toxic polyneuropathies
Neuropathy	R43	Disturbances of smell and taste
Other	G93	Other disorders of the brain
Psychiatric	F29	Unspecified psychosis not due to a substance or known physiological condition
Seizure	G40	Epilepsy and recurrent seizures
Vascular	G45	Transient cerebral ischemic attacks and related syndromes
Vascular	G46	Vascular syndromes of brain in cerebrovascular disease
Vascular	I60	Nontraumatic subarachnoid hemorrhage
Vascular	I61	Nontraumatic intracerebral hemorrhage
Vascular	I62	Other and unspecified nontraumatic intracranial hemorrhage
Vascular	I67	Other cerebrovascular diseases
Vision	H54	Blindness and low vision

Table 3. Mapping of neurological disease categories to ICD-10 category codes and their descriptions. [a]To standardize the collection of ICD codes across diverse contributing healthcare systems and to mitigate coding discrepancies, we used ICD codes at the categorical level (*e.g.*, the first 3 alphanumeric characters before the decimal point for ICD-10).

Data availability

The 4CE consortium does not have permission from each individual contributing healthcare system to release electronic health records data for public access. Only aggregate data were shared by healthcare systems for this study. All aggregate data in a de-identified fashion are available for download (http://www.covidclinical.net).

Code availability

Our analysis results are available in browsable R notebooks (https://github.com/covidclinical/Phase1.1NeuroRCode/) under CC BY 4.0, with source code distributed under a BSD 3-Clause License. Figure 1 was created with Biorender (Biorender.com). Figures 2, 3, 4 as well as all supplementary figures were created using R v.4.0.3 (http://www.R-project.org)[45] and the package ggplot2[46].

References

1. Mahase, E. Covid-19: WHO declares pandemic because of "alarming levels" of spread, severity, and inaction. *BMJ* **368**, m1036 (2020).
2. Koralnik, I. J. & Tyler, K. L. COVID-19: A global threat to the nervous system. *Ann. Neurol.* **88**, 1–11 (2020).
3. Pezzini, A. & Padovani, A. Lifting the mask on neurological manifestations of COVID-19. *Nat. Rev. Neurol.* **16**, 636–644 (2020).
4. Eskandar, E. N. *et al.* Neurologic syndromes predict higher in-hospital mortality in COVID-19. *Neurology* **96**, e1527–e1538 (2021).
5. Ellul, M. A. *et al.* Neurological associations of COVID-19. *Lancet Neurol.* **19**, 767–783 (2020).
6. Moriguchi, T. *et al.* A first case of meningitis/encephalitis associated with SARS-Coronavirus-2. *Int. J. Infect. Dis.* **94**, 55–58 (2020).
7. Novi, G. *et al.* Acute disseminated encephalomyelitis after SARS-CoV-2 infection. *Neurol. Neuroimmunol. Neuroinflamm.* **7**, e797 (2020).
8. Dinkin, M. *et al.* COVID-19 presenting with ophthalmoparesis from cranial nerve palsy. *Neurology* **95**, 221–223 (2020).
9. Toscano, G. *et al.* Guillain-Barré syndrome associated with SARS-CoV-2. *N. Engl. J. Med.* **382**, 2574–2576 (2020).
10. Dalakas, M. C. Guillain-Barré syndrome: The first documented COVID-19-triggered autoimmune neurologic disease: More to come with myositis in the offing. *Neurol. Neuroimmunol. Neuroinflamm.* **7**, e781 (2020).
11. Fragiel, M. *et al.* Incidence, clinical, risk factors and outcomes of Guillain-Barré in Covid-19. *Ann. Neurol.* **89**, 598–603 (2021).
12. Paterson, R. W. *et al.* The emerging spectrum of COVID-19 neurology: Clinical, radiological and laboratory findings. *Brain* **143**, awaa240 (2020).
13. Spinato, G. *et al.* Alterations in smell or taste in mildly symptomatic outpatients with SARS-CoV-2 infection. *JAMA* **323**, 2089 (2020).
14. Xydakis, M. S. *et al.* Post-viral effects of COVID-19 in the olfactory system and their implications. *Lancet Neurol.* https://doi.org/10.1016/s1474-4422(21)00182-4 (2021).
15. Rogers, J. P. *et al.* Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: A systematic review and meta-analysis with comparison to the COVID-19 pandemic. *Lancet Psychiatry* **7**, 611–627 (2020).
16. Oxley, T. J. *et al.* Large-vessel stroke as a presenting feature of Covid-19 in the young. *N. Engl. J. Med.* **382**, e60 (2020).

17. Abdel-Mannan, O. *et al.* Neurologic and radiographic findings associated with COVID-19 infection in children. *JAMA Neurol.* https://doi.org/10.1001/jamaneurol.2020.2687 (2020).
18. Ray, S. T. J. *et al.* Neurological manifestations of SARS-CoV-2 infection in hospitalised children and adolescents in the UK: A prospective national cohort study. *Lancet Child Adolesc. Health* https://doi.org/10.1016/s2352-4642(21)00193-0 (2021).
19. Mao, L. *et al.* Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. *JAMA Neurol.* **77**, 683 (2020).
20. Varatharaj, A. *et al.* Neurological and neuropsychiatric complications of COVID-19 in 153 patients: A UK-wide surveillance study. *Lancet Psychiatry* **7**, 875–882 (2020).
21. Drake, T. M. *et al.* Characterisation of in-hospital complications associated with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol UK: A prospective, multicentre cohort study. *Lancet* **398**, 223–237 (2021).
22. Benussi, A. *et al.* Clinical characteristics and outcomes of inpatients with neurologic disease and COVID-19 in Brescia, Lombardy, Italy. *Neurology* **95**, e910–e920 (2020).
23. Helms, J. *et al.* Neurologic features in severe SARS-CoV-2 infection. *N. Engl. J. Med.* **382**, 2268–2270 (2020).
24. Liotta, E. M. *et al.* Frequent neurologic manifestations and encephalopathy-associated morbidity in Covid-19 patients. *Ann. Clin. Transl. Neurol.* **324**, 603 (2020).
25. Frontera, J. A. *et al.* A prospective study of neurologic disorders in hospitalized COVID-19 patients in New York City. *Neurology* https://doi.org/10.1212/WNL.0000000000010979 (2020).
26. Taquet, M., Geddes, J. R., Husain, M., Luciano, S. & Harrison, P. J. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: A retrospective cohort study using electronic health records. *Lancet Psychiatry* **8**, 416–427 (2021).
27. Chou, S.H.-Y. *et al.* Global incidence of neurological manifestations among patients hospitalized with COVID-19—A report for the GCS-NeuroCOVID consortium and the ENERGY consortium. *JAMA Netw. Open* **4**, e2112131 (2021).
28. Brat, G. A. *et al.* International electronic health record-derived COVID-19 clinical course profiles: The 4CE consortium. *NPJ Digit. Med.* **3**, 109 (2020).
29. Weber, G. M. *et al.* International comparisons of harmonized laboratory value trajectories to predict severe COVID-19: Leveraging the 4CE collaborative across 342 hospitals and 6 countries: A retrospective cohort study. *medRxiv* https://doi.org/10.1101/2020.12.16.20247684 (2020).
30. Klann, J. G. *et al.* Validation of an internationally derived patient severity phenotype to support COVID-19 analytics from electronic health record data. *J. Am. Med. Inform. Assoc.* **28**, 1411–1420. https://doi.org/10.1093/jamia/ocab018 (2021).
31. Bourgeois, F. T. *et al.* International analysis of electronic health records of children and youth hospitalized with COVID-19 infection in 6 countries. *Jama Netw. Open* **4**, e2112596 (2021).
32. Pun, B. T. *et al.* Prevalence and risk factors for delirium in critically ill patients with COVID-19 (COVID-D): A multicentre cohort study. *Lancet Respir. Med.* **9**, 239–250 (2021).
33. Cosentino, G. *et al.* Neuropathological findings from COVID-19 patients with neurological symptoms argue against a direct brain invasion of SARS-CoV-2: A critical systematic review. *Eur. J. Neurol.* https://doi.org/10.1111/ene.15045 (2021).
34. Zubair, A. S. *et al.* Neuropathogenesis and neurologic manifestations of the coronaviruses in the age of coronavirus disease 2019: A review. *JAMA Neurol.* **77**, 1018–1027 (2020).
35. Thakur, K. T. *et al.* COVID-19 neuropathology at Columbia University Irving Medical Center/New York Presbyterian Hospital. *Brain* https://doi.org/10.1093/brain/awab148 (2021).
36. Siegler, J. E. *et al.* Cerebrovascular events and outcomes in hospitalized patients with COVID-19: The SVIN COVID-19 Multinational Registry. *Int. J. Stroke* **52**, 1747493020959216 (2020).
37. Levi, M., Thachil, J., Iba, T. & Levy, J. H. Coagulation abnormalities and thrombosis in patients with COVID-19. *Lancet. Haematol.* **7**, e438–e440 (2020).
38. Guan, W.-J. *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
39. Bagnato, S. *et al.* Critical illness myopathy after COVID-19. *Int. J. Infect. Dis.* **99**, 276–278 (2020).
40. McNett, M. *et al.* The Global Consortium Study of Neurological Dysfunction in COVID-19 (GCS-NeuroCOVID): Development of case report forms for global use. *Neurocrit. Care* **33**, 793–828 (2020).
41. Zhang, Y. *et al.* High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). *Nat. Protoc.* **14**, 3426–3444 (2019).
42. Wasserman, L. *All of Statistics: A Concise Course in Statistical Inference* (Springer, 2013).
43. Fisher, R. A. On the interpretation of χ^2 from contingency tables, and the calculation of P. *J. R. Stat. Soc.* **85**, 87–94 (1922).
44. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125**, 279–284 (2001).
45. Team, R. C. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).
46. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).

Author contributions
T.L., A.G.S., J.S., G.O., P.A., S.V., D.M., and Z.X. contributed to concept and design of the study. B.B., N.L., Y.L., M.M., K.N., L.P., M.S., E.S., A.T., I.K., S.V., D.M., and Z.X. contributed to data collection. T.L., A.G.S., J.S., C.H., K.N., T.C., P.A., S.V., D.M., and Z.X. contributed to data analysis. T.L., A.G.S., J.S., C.H., A.S., B.B., N.L., Y.L., J.M., L.P., T.C., G.O., P.A., S.V., D.M., and Z.X. contributed to drafting and editing the manuscript. All named authors have approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The Consortium for Clinical Characterization of COVID-19 by EHR (4CE)

James R. Aaron[12], Giuseppe Agapito[13], Adem Albayrak[14], Mario Alessiani[15], Danilo F. Amendola[16], François Angoulvant[17], Li L. L. J. Anthony[18], Bruce J. Aronow[19], Andrew Atz[20], James Balshi[21], Douglas S. Bell[22], Antonio Bellasi[23], Riccardo Bellazzi[24], Vincent Benoit[25], Michele Beraghi[26], José Luis Bernal Sobrino[27], Mélodie Bernaux[28], Romain Bey[25], Alvar Blanco Martínez[27], Martin Boeker[29], Clara-Lea Bonzel[2], John Booth[30], Silvano Bosari[31], Florence T. Bourgeois[2], Robert L. Bradford[32], Gabriel A. Brat[2], Stéphane Bréant[25], Nicholas W. Brown[2], William A. Bryant[30], Mauro Bucalo[33], Anita Burgun[34], Mario Cannataro[35], Aldo Carmona[36], Charlotte Caucheteux[37], Julien Champ[38], Krista Chen[39], Jin Chen[40], Luca Chiovato[41,42], Lorenzo Chiudinelli[23], James J. Cimino[43], Tiago K. Colicchio[43], Sylvie Cormont[25], Sébastien Cossin[44], Jean B. Craig[45], Juan Luis Cruz Bermúdez[27], Jaime Cruz Rojo[27], Arianna Dagliati[24], Mohamad Daniar[46], Christel Daniel[25], Anahita Davoudi[1], Batsal Devkota[47], Julien Dubiel[25], Loic Esteve[48], Shirley Fan[49], Robert W. Follett[22], Paula S. A. Gaiolla[50], Thomas Ganslandt[51], Noelia García Barrio[27], Lana X. Garmire[11], Nils Gehlenborg[2], Alon Geva[39,52], Tobias Gradinger[51], Alexandre Gramfort[37], Romain Griffier[44], Nicolas Griffon[25], Olivier Grisel[37], David A. Hanauer[53], Christian Haverkamp[29], Bing He[11], Darren W. Henderson[12], Martin Hilka[25], John H. Holmes[1], Petar Horki[29], Kenneth M. Huling[2], Meghan R. Hutch[6], Richard W. Issitt[30], Anne Sophie Jannot[34], Vianney Jouhet[44], Ramakanth Kavuluru[12], Mark S. Keller[2], Katie Kirchoff[54], Jeffrey G. Klann[55], Ian D. Krantz[1,56], Detlef Kraska[57], Ashok K. Krishnamurthy[58], Sehi L'Yi[2], Judith Leblanc[59], Andressa R. R. Leite[16], Guillaume Lemaitre[37], Leslie Lenert[45], Damien Leprovost[60], Molei Liu[61], Sarah Lozano-Zahonero[9], Kristine E. Lynch[62], Sadiqa Mahmood[14], Sarah Maidlow[63], Adeline C. Makoudjou Tchendjou[29], Alberto Malovini[64], Kenneth D. Mandl[39], Chengsheng Mao[6], Anupama Maram[65], Patricia Martel[66], Aaron J. Masino[67], Michael E. Matheny[68], Thomas Maulhardt[29], Maria Mazzitelli[35], Michael T. McDuffie[2], Arthur Mensch[69], Fatima Ashraf[70], Marianna Milano[35], Marcos F. Minicucci[50], Bertrand Moal[44], Cinta Moraleda[71], Jeffrey S. Morris[1], Karyn L. Moshal[72], Sajad Mousavi[2], Douglas A. Murad[22], Shawn N. Murphy[73], Thomas P. Naughton[65], Antoine Neuraz[34], James B. Norman[2], Jihad Obeid[45], Marina P. Okoshi[74], Karen L. Olson[39,75], Nina Orlova[25], Brian D. Ostasiewski[76], Nathan P. Palmer[2], Nicolas Paris[25], Miguel Pedrera Jimenez[27], Emily R. Pfaff[32], Danielle Pillion[2], Hans U. Prokosch[77], Robson A. Prudente[16], Víctor Quirós González[27], Rachel B. Ramoni[78], Maryna Raskin[14], Siegbert Rieg[29], Gustavo Roig Domínguez[27], Pablo Rojo[27], Carlos Sáez[79], Elisa Salamanca[25], Arnaud Sandrin[25], Janaina C. C. Santos[16,80],

Maria Savino[81], Juergen Schuettler[82], Luigia Scudeller[31], Neil J. Sebire[30],
Pablo Serrano Balazote[27], Patricia Serre[25], Arnaud Serret-Larmande[2], Zahra Shakeri[2],
Domenick Silvio[63], Piotr Sliz[39], Charles Sonday[83], Anastasia Spiridou[30], Bryce W. Q. Tan[84],
Byorn W. L. Tan[84], Suzana E. Tanni[74], Deanne M. Taylor[1,47], Ana I. Terriza-Torres[27],
Valentina Tibollo[64], Patric Tippmann[29], Carlo Torti[35], Enrico M. Trecarichi[35], Yi-Ju Tseng[85],
Andrew K. Vallejos[86], Gael Varoquaux[37,87], Margaret Vella[2], Jill-Jênn Vie[88], Michele Vitacca[89],
Kavishwar B. Wagholikar[55], Lemuel R. Waitman[90], Demian Wassermann[37], Griffin M. Weber[2],
Yuan William[2], Nadir Yehya[67], Alberto Zambelli[91], Harrison G. Zhang[2], Daniela Zoeller[29] &
Chiara Zucco[35]

[12]Department of Biomedical Informatics, University of Kentucky, Lexington, KY, USA. [13]Department of Legal, Economic and Social Sciences, University Magna Graecia of Catanzaro, Catanzaro, Italy. [14]Health Catalyst, INC., Cambridge, MA, USA. [15]Department of Surgery, ASST Pavia, Lombardia Region Health System, Pavia, Italy. [16]Clinical Research Unit of Botucatu Medical School, São Paulo State University, Botucatu, Brazil. [17]Pediatric Emergency Department, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris, Paris, France. [18]National Center for Infectious Diseases, Tan Tock Seng Hospital, Singapore, Singapore. [19]Departments of Biomedical Informatics, Pediatrics, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, OH, USA. [20]Department of Pediatrics, Medical University of South Carolina, Charleston, SC, USA. [21]Department of Surgery, St. Luke's University Health Network, Bethlehem, PA, USA. [22]Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. [23]UOC Ricerca, Innovazione e Brand Reputation, ASST Papa Giovanni XXIII, Bergamo, Italy. [24]Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy. [25]IT Department, Innovation & Data, APHP Greater Paris University Hospital, Paris, France. [26]I.T. Department, ASST Pavia, Pavia, Italy. [27]Health Informatics, Hospital Universitario, 12 de Octubre, Madrid, Spain. [28]Strategy and Transformation Department, APHP Greater Paris University Hospital, Paris, France. [29]Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany. [30]Digital Research, Informatics and Virtual Environments (DRIVE), Great Ormond Street Hospital for Children, London, UK. [31]Scientific Direction, IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano, Milan, Italy. [32]North Carolina Translational and Clinical Sciences (NC TraCS) Institute, University of North Carolina, Chapel Hill, NC, USA. [33]BIOMERIS (BIOMedical Research Informatics Solutions), Pavia, Italy. [34]Department of Biomedical Informatics, HEGP, APHP Greater Paris University Hospital, Paris, France. [35]Department of Medical and Surgical Sciences, Data Analytics Research Center, University Magna Graecia of Catanzaro, Catanzaro, Italy. [36]Department of Anesthesia, St. Luke's University Health Network, Bethlehem, PA, USA. [37]Université Paris-Saclay, Inria, CEA, Palaiseau, France. [38]INRIA Sophia-Antipolis–ZENITH Team, LIRMM, Montpellier, France. [39]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA. [40]Department of Internal Medicine, University of Kentucky, Lexington, KY, USA. [41]Unit of Internal Medicine and Endocrinology, Istituti Clinici Scientifici Maugeri SpA SB IRCCS, Pavia, Italy. [42]Department of Internal Medicine and Therapeutics, University of Pavia, Pavia, Italy. [43]Informatics Institute, University of Alabama at Birmingham, Birmingham, AL, USA. [44]IAM Unit, Bordeaux University Hospital/ERIAS-Inserm U1219 BPH, Bordeaux, France. [45]Biomedical Informatics Center, Medical University of South Carolina, Charleston, SC, USA. [46]Clinical Research Informatics, Boston Children's Hospital, Boston, MA, USA. [47]Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [48]SED/SIERRA, Inria Centre de Paris, Paris, France. [49]Health Information Technology & Services, University of Michigan, Ann Arbor, MI, USA. [50]Internal Medicine Department, Botucatu Medical School, São Paulo State University, Botucatu, Brazil. [51]Heinrich-Lanz-Center for Digital Health, University Medicine Mannheim, Heidelberg University, Mannheim, Germany. [52]Department of Anesthesiology, Critical Care, and Pain Medicine Boston Children's Hospital, Boston, MA, USA. [53]Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, MI, USA. [54]MSHI Medical University of South Carolina, Charleston, SC, USA. [55]Department of Medicine, Massachusetts General Hospital, Boston, MA, USA. [56]Division of Human Genetics, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. [57]Center for Medical Information and Communication Technology, University Hospital, Erlangen, Germany. [58]Renaissance Computing Institute/Department of Computer Science, University of North Carolina, Chapel Hill, NC, USA. [59]Clinical Research Unit, Saint Antoine Hospital, APHP Greater Paris University Hospital, Paris, France. [60]Clevy.io, Paris, France. [61]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [62]VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, USA. [63]MICHR Informatics, University of Michigan, Ann Arbor, MI, USA. [64]Laboratory of Informatics and Systems Engineering for Clinical Research, Istituti Clinici Scientifici Maugeri SpA SB IRCCS, Pavia, Italy. [65]Harvard Catalyst, Harvard Medical School, Boston, MA, USA. [66]Clinical Research Unit, Paris Saclay, APHP Greater Paris University Hospital, Paris, France. [67]Department of Anesthesiology and Critical Care, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [68]VA Informatics and Computing Infrastructure, Tennessee Valley Healthcare System, Veterans Affairs Medical Center, Washington, D.C., USA. [69]École Normale Supérieure, PSL Université Paris, Paris, France. [70]BIG-ARC, The University of Texas Health Science Center at Houston, School of Biomedical Informatics, Houston, TX, USA. [71]Pediatric Infectious Disease Department, Hospital Universitario, 12 de Octubre, Madrid, Spain. [72]Department of Infectious Diseases, Great Ormond Street Hospital for Children, London, UK. [73]Department of Neurology, Massachusetts General Hospital, Boston, MA, USA. [74]Internal Medicine Department of Botucatu Medical School, São Paulo State University, Botucatu, Brazil. [75]Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA. [76]Center for Biomedical Informatics, Wake Forest School of Medicine, Winston-Salem, NC, USA. [77]Department of Medical Informatics, University of Erlangen-Nürnberg, Erlangen, Germany. [78]Department of Veterans Affairs, Office of Research and Development, Baltimore, MD, USA. [79]Biomedical Data Science Lab, ITACA Institute, Universitat

Politècnica de València, València, Spain. [80]Nurse Department of FMB-Medicine School of Botucatu, Botucatu, Brazil. [81]Management Engineering, ASST Pavia, Lombardia Region Health System, Pavia, Italy. [82]Department of Anesthesiology, University Hospital Erlangen, FAU Erlangen-Nürnberg, Erlangen, Germany. [83]Critical Care Medicine, Department of Medicine, St. Luke's University Health Network, Bethlehem, PA, USA. [84]Department of Medicine, National University Hospital, Singapore, Singapore. [85]Department of Information Management, Chang Gung University, Taoyuan, Taiwan. [86]Clinical & Translational Science Institute, Medical College of Wisconsin, Milwaukee, WI, USA. [87]Montréal Neurological Institute, McGill University, Montreal, Canada. [88]SequeL, Inria Lille, Villeneuve-d'Ascq, France. [89]Respiratory Department, ICS S Maugeri IRCCS, Pavia, Italy. [90]Department of Health Management and Informatics, University of Missouri, Columbia, MO, USA. [91]Department of Oncology, ASST Papa Giovanni XXIII, Bergamo, Italy.