

# Multinational Prevalence of Neurological Phenotypes in Patients Hospitalized with COVID-19

Trang T. Le, PhD<sup>1\*</sup>, Alba Gutiérrez-Sacristán, PhD<sup>2\*</sup>, Jiyeon Son, MD<sup>3\*</sup>, Chuan Hong, PhD<sup>3</sup>, Andrew M. South, MD, MS<sup>4</sup>, Brett K. Beaulieu-Jones, PhD<sup>3</sup>, Ne Hooi Will Loh, MBBS<sup>5</sup>, Yuan Luo, PhD<sup>6</sup>, Michele Morris, BA<sup>7</sup>, Kee Yuan Ngiam, MBBS, MRCS, MMed<sup>8</sup>, Lav P. Patel, MS<sup>9</sup>, Malarkodi J. Samayamuthu, MD<sup>7</sup>, Emily Schriver, MS<sup>10</sup>, Amelia LM Tan, PhD<sup>2</sup>, Jason Moore, PhD<sup>1</sup>, Tianxi Cai, ScD<sup>2</sup>, Gilbert S. Omenn, MD, PhD<sup>11</sup>, Paul Avillach, MD, PhD<sup>2</sup>, Isaac S. Kohane, MD, PhD<sup>2</sup>, **4CE Consortium**, Shyam Visweswaran, MD, PhD<sup>7\*\*</sup>, Danielle L. Mowery, PhD<sup>1\*\*</sup>, Zongqi Xia, MD, PhD<sup>12\*\*</sup>

\* Drs. Le, Gutiérrez-Sacristán and Son share co-first authorship.

\*\* Drs. Visweswaran, Mowery and Xia share co-senior authorship.

4CE Consortium: International Consortium for Clinical Characterization of COVID-19 by EHR

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, Pittsburgh, PA, USA
4. Department of Pediatrics, Wake Forest School of Medicine, Winston Salem, NC, USA
5. Department of Critical Care, National University Health Systems, 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

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
12. Department of Neurology, University of Pittsburgh, PA, USA

**Count:**

Title (without subtitle):	10 words, 91 characters (with spaces)
Abstract:	244 words
Manuscript:	3,100 words (Introduction: 420 words; Discussion: 1011 words)
References:	38
Tables:	3
Figures:	4 (4 color figures)
Supplementary Tables:	2
Supplementary Figures:	5

**Running Title:** Prevalence of Neurological Phenotypes in COVID-19

**Key Words:** COVID-19, neurology, electronic health records, public health

**Corresponding Authors:**

Zongqi Xia, MD, PhD

Department of Neurology, University of Pittsburgh

Biomedical Science Tower 3, Suite 7014, 3501 5<sup>th</sup> Avenue, Pittsburgh, PA 15260

Phone: 412-383-5377; Fax: 412-648-7233; Email: [zxia1@post.harvard.edu](mailto:zxia1@post.harvard.edu)

## Abstract

**OBJECTIVE:** Neurological complications can worsen outcomes in COVID-19. We defined the prevalence of a wide range of neurological conditions among patients hospitalized with COVID-19 in geographically diverse multinational populations.

**METHODS:** Using electronic health record (EHR) data from 348 participating hospitals across 6 countries and 3 continents between January and September 2020, we performed a cross-sectional study of hospitalized adult and pediatric patients with a positive SARS-CoV-2 reverse transcription polymerase chain reaction test, both with and without severe COVID-19. We assessed the frequency of each disease category and 3-character International Classification of Disease (ICD) code of neurological diseases by countries, sites, time before and after admission for COVID-19, and COVID-19 severity.

**RESULTS:** Among the 35,177 hospitalized patients with SARS-CoV-2 infection, there was increased prevalence of disorders of consciousness (5.8%, 95% confidence interval [CI]: 3.7%-7.8%,  $p_{\text{FDR}} < .001$ ) and unspecified disorders of the brain (8.1%, 95%CI: 5.7%-10.5%,  $p_{\text{FDR}} < .001$ ), compared to pre-admission prevalence. During hospitalization, patients who experienced severe COVID-19 status had 22% (95%CI: 19%-25%) increase in the relative risk (RR) of disorders of consciousness, 24% (95%CI: 13%-35%) increase in other cerebrovascular diseases, 34% (95%CI: 20%-50%) increase in nontraumatic intracranial hemorrhage, 37% (95%CI: 17%-60%) increase in encephalitis and/or myelitis, and 72% (95%CI: 67%-77%) increase in myopathy compared to those who never experienced severe disease.

**INTERPRETATION:** Using an international network and common EHR data elements, we highlight an increase in the prevalence of central and peripheral neurological phenotypes in

patients hospitalized with SARS-CoV-2 infection, particularly among those with severe disease.

## Introduction

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.<sup>1</sup> Increasing evidence points to the multi-organ involvement of COVID-19, including the nervous system, which increases morbidity and mortality.

Understanding the neurological manifestations is essential for developing strategies to mitigate disability and death from the neurological consequences of COVID-19.

Many case reports and reviews have highlighted neurological phenotypes in adults with COVID-19, including cerebrovascular disease<sup>2</sup>, meningoencephalitis and encephalomyelitis<sup>3,4</sup>, encephalopathy<sup>5</sup>, headache<sup>6</sup>, cranial neuropathies<sup>7</sup>, Guillain-Barré syndrome<sup>8,9</sup>, plexopathy<sup>10</sup>, anosmia and ageusia<sup>11,12</sup>, and cognitive dysfunction<sup>13</sup>. Children with COVID-19 have similar presentations, including ischemic stroke, encephalopathy, headache, and muscle weakness<sup>14,15</sup>. Thus, neurological complications in the setting of COVID-19 affect both the central and the peripheral nervous system. Prior studies reported the prevalence of COVID-19-associated neurological conditions at country level, including China<sup>16</sup>, the United Kingdom<sup>17</sup>, and Italy<sup>18</sup>, but few studies have used validated common data elements to examine the prevalence of neurologic conditions across countries. While studies have described the spectrum of neurological conditions associated with severe COVID-19<sup>19–21</sup>, no multinational study has compared the prevalence of neurological complications in patients with severe versus non-severe disease, based on respiratory and/or critical illness status. Given the health consequences of neurological complications, recognizing the associated neurological complications could inform prevention, diagnosis, and treatment.

Clinical data in electronic health records (EHRs) enable studies across institutions to estimate the prevalence of neurological conditions in patients with COVID-19. There are several challenges when using international and multi-institutional EHR data for clinical discovery: (1) adopting common data elements and standardization processes for representing and aggregating clinical events, (2) sharing clinical data across institutions 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), and (3) rapidly developing and deploying analytical solutions at scale by leveraging existing informatics infrastructures and frameworks. Our team created the International Consortium for Clinical Characterization of COVID-19 by EHR (4CE; [www.covidclinical.net](http://www.covidclinical.net)), standardizing and aggregating EHR data from nearly 350 hospitals across six countries to identify and address critical clinical and epidemiological questions relevant to COVID-19.<sup>22,23</sup> Leveraging the highly scalable, multinational networks of the 4CE consortium, we estimated the prevalence of neurological conditions in hospitalized patients with reverse transcription polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection by site and country, and compared the differences in the prevalence of neurological conditions between patients with and those without severe disease based on the internationally validated 4CE criteria.<sup>24</sup>

## Methods

### *Patients and Data*

4CE contributing sites began in March 2020 to collect EHR data from hospitalized patients with positive SARS-CoV-2 RT-PCR tests. The analyzed data spanned 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 sites to protect site confidentiality. The institutional review board of each participating site approved the sharing of anonymous, aggregate data in compliance with multi-national patient privacy laws exempting the requirement for individual patient consent. A small level of obfuscation to preserve site-specific privacy and to reduce the risk of patient re-identification may slightly but not significantly affect the total patient counts.

Using the standardized 4CE common data model<sup>22,23</sup>, we collected demographics (age, sex, 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, defined according to the 4CE definition.<sup>24</sup> Of the 45 contributing sites, five sites (all from Italy) provided only ICD-9 codes while the remaining 40 sites provided predominantly ICD-10 codes. As such, we used ICD-10 data for the main analyses and ICD-9 data for supplementary

analyses. 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”).

For all patients hospitalized with COVID-19, we collected ICD codes at two time periods, *before* and *after* the admission date. The period before COVID-19 hospitalization ranged from -365 days to -15 days preceding the admission date. Inclusion of EHR data up to one year before admission is a pragmatic decision to balance available data from all sites 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 two weeks preceding the 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<sup>24</sup> (**Fig. 1**).

## *Exposures*

We first examined all hospitalized patients with positive SARS-CoV-2 PCR tests. We then examined hospitalized patients with COVID-19 who met any of the definitions for severe COVID-19<sup>24</sup>, including advanced respiratory care management at any point during their

hospitalization. Definitions of severe COVID-19 included diagnoses, procedures, laboratory results, or medications (**Table 1**).

## *Neurological Outcomes*

We queried 21 ICD-10 codes in 12 categories pertaining to neurological phenotypes following a 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 2**).

## *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 site (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 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 site 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.<sup>25</sup> Finally, we computed the  $p$ -values using Fisher's exact test<sup>26</sup> and corrected for multiple hypothesis testing with Benjamini–Hochberg's false discovery rate (FDR) procedure.<sup>27</sup> A result was statistically significant if  $p_{\text{FDR}} < 0.05$ .

### *Data / Code Availability*

The 4CE consortium does not have permission from each individual contributing site to release EHR data for public access. 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.

## Results

### *Demographics*

Among the 35,177 hospitalized patients with PCR-confirmed SARS-CoV-2 infection, aggregate demographic data were available for 34,647 patients (98.5%) from 348 hospitals (affiliated with 43 sites, after excluding two US sites with incomplete demographic data). 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 sex (**Fig. 2**). Variation in the proportions of severe COVID-19 cases among the US sites was large, with 18 sites reporting >90% severe cases. The proportions of severe COVID-19 cases were higher in France than Germany, Italy, Spain, and Singapore. There was no clear relationship between COVID-19 severity and median age (**Fig. 2B**). Most sites in Europe did not report race. Among the US sites, there was a disproportionately high proportion of 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**).

### Prevalence of Neurological Conditions in All Patients

Using the prevalence before admission as reference, we found that the majority of contributing sites (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%,  $p_{\text{FDR}} < 0.001$ ) after admission (**Fig. 3, eFig. 1**). Similarly, 84% of sites reported an increase in the proportion of

patients with “Unspecified disorders of the brain” (ICD-10 G93, including “encephalopathy”) with a mean increase of 8.1% (5.7-10.5%,  $p_{\text{FDR}} < 0.001$ ) after admission. (see [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.

## Prevalence of Neurological Conditions in Severe COVID-19

For each neurological ICD code, we used Fisher’s exact test to examine the enrichment or depletion of the condition among patients who ever experienced severe disease when compared to those who never experienced severe disease (**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 reporting R41 (disorder of consciousness) was  $2^{0.283}$  or  $\sim 1.22$  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%). We presented in **Table 3** the statistically significant associations of neurological conditions and severe COVID-19 status ( $p_{\text{FDR}} < 0.05$ ). Please refer to the online [interactive data table](#) or the [results](#) directory of the project online

repository for the full table containing the LOE, 95% confidence intervals, and  $p$  values for all neurological ICD-codes.

In the period after hospital admission date, using the patients who never experienced severe disease as reference, we observed a significantly higher proportion of patients with severe disease who had “other symptoms and signs involving cognitive functions and awareness” (R41:  $RRD_{after}=22\%$ ), “other cerebrovascular 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”:  $RRD_{after}=72\%$ ). In contrast, we noted a significantly lower proportion of patients with severe disease who 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 (**Table 3, Fig. 4**). Subgroup analysis comparing the US and non-US sites yielded mostly consistent and expected results (**eFig. 2**).

## Exploratory Analysis

As 11 sites in the US and Italy contributed data entirely or partially consisting of ICD-9 codes, we performed separate subgroup analyses using only the ICD-9 data (**eTable 1**), given that one-to-one mapping of some ICD-9 codes to ICD-10 codes was not feasible. To summarize, there was increased prevalence of “disorders of consciousness” and “other neurological conditions” in patients after admission date when compared to before admission date, similar to the ICD-10 data analysis (**eFig. 3**, **eFig. 4**). The increase in these conditions appear to be driven by the US sites but not the Italian sites. There was no statistically significant difference when examining the change in prevalence of individual 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 notable differences involving seizure and cerebrovascular events that would require caution in interpretation due to differences in sample size (see **Supplementary Material**, **eFig. 5**, **eTable 2**).

## Discussion

In this study, we implemented a research strategy for rapidly aggregating EHR-derived clinical facts across multiple international sites while preserving patient privacy. We report the prevalence of central and peripheral neurological conditions among patients hospitalized with SARS-CoV-2 infection across geographically diverse hospital systems from six countries, using the time period before hospitalization as reference. Importantly, we further compared the prevalence of neurological conditions between patients with and without severe COVID-19 disease.

We found disorders of consciousness and other disorders of the brain as the most prevalent neurological phenotypes (based on ICD codes) among all patients hospitalized with positive SARS-CoV-2, consistent with prior reports.<sup>16,17,21</sup> Interestingly, our study showed low prevalence of early COVID-19 symptoms such as alterations in smell and taste<sup>12</sup>, likely attributable to the inconsistent documentation of these symptoms in the EHR for the hospitalized patient population, particularly those with severe COVID-19. Further, we did not find statistically significant changes in the prevalence of other previously reported neurological conditions such as dizziness, encephalopathy, headache, seizure, vascular, and vision disorders after admission to the hospital. These discrepancies may be partially due to methodological differences, as our analysis examined the change in prevalence using the data before admission as the reference.

Our major finding indicates that a significantly higher proportion of hospitalized patients with severe COVID-19 had disorders of consciousness, encephalitis and/or myelitis,

cerebrovascular events, and myopathy when compared to patients who never had severe disease. While these findings are largely consistent with prior reports<sup>16,20,21</sup>, our contribution is the report of findings from geographically diverse multinational sites. First among these findings, disorders of consciousness, including altered mental status and encephalopathy as well as encephalitis, are consistently reported neurological complications of COVID-19. Sedation for advanced respiratory support and hypoxemia from respiratory failure may partially explain the association of disorders of consciousness with severe COVID-19. Notably, encephalopathy is independently associated with high morbidity and mortality in COVID-19.<sup>21</sup> Future studies will be important to identify whether the mechanisms underlying encephalitis (with or without myelitis) are direct central nervous system (CNS) invasion by SARS-CoV-2, acute systemic inflammation with secondary CNS involvement, and/or post-infectious immune-mediated effect on the CNS.<sup>28,29</sup>

Second, 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 (38% for ischemic stroke and 58% for intracranial hemorrhage)<sup>30</sup> and morbidity. COVID-19 might increase the risk of ischemic stroke through diverse 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<sup>31</sup>. 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<sup>32,33</sup>. Patients with severe COVID-19 may have additional risk factors such as hypertension or cardiovascular disease that could further drive cerebrovascular diseases.<sup>34</sup>

Finally, while myopathy is common among COVID-19 patients, its mechanism remains unclear. Given that SARS-CoV-2 interacts with the spike domain of ACE2 on host cells<sup>35</sup> and skeletal muscle tissues express ACE2<sup>36</sup>, muscle may be susceptible to direct SARS-CoV-2 infection. However, myopathy in patients with severe COVID-19 is likely attributable to prolonged or severe critical illness<sup>37</sup>, complications due to multi-organ involvement, or medication-induced myotoxicity (e.g., Hydroxychloroquine, steroids). Our current study design cannot differentiate whether the neurological phenotypes 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 objective evaluation or proper documentation and coding for these conditions.

Our study has limitations as the result of trade-offs to standardize data collection from multinational sites while strictly preserving patient privacy and adhering to multinational privacy laws across all contributing sites. 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 collection of ICD codes across contributing sites 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). Thus, further characterization of certain conditions such as “other disorders of the brain” was not feasible at this stage. Second, the proportions of severe cases varied across sites. Adding the number of patients from all sites for each ICD code retained statistical power in the severity analysis, though data from the larger sites likely drove the findings. Third, because we aggregated data across sites, we were unable to consolidate all related ICD codes (e.g., organizing into PheCode<sup>38</sup>) at the individual patient level. The 4CE consortium is addressing these issues to prepare for the next phase of the analyses by developing a framework for conducting patient-level analyses at each site. Fourth, we might not have captured all pre-admission EHR data if patients did not receive all of their care in the same hospital system as the COVID-19 admission. This is a limitation common to all research using EHR data from non-universal health systems. Finally, contributing sites with small patient counts used obfuscation to reduce the risk of re-identifying sites, but the effect of obfuscation is negligible because few sites 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 replicated prior reports from well-characterized but often smaller and single-center cohort studies.

We report the first multinational prevalence study of a broad spectrum of central and peripheral neurological phenotypes in hospitalized patients with PCR-confirmed SARS-CoV-2 infection, particularly those with severe disease, using a common EHR data collection method. Our EHR-based efforts 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 neurological phenotypes, and examine long-term neurological sequelae in patients with COVID-19.

## Acknowledgement

We greatly appreciate Margaret Vella in her coordination of the 4CE consortium effort and assistance in the manuscript submission.

## Author Contribution

TL, AGS, JS, GO, PA, SV, DM, and ZX contributed to concept and design of the study. BB, NL, YL, MM, KN, LP, MS, ES, AT, IK, SV, DM, and ZX contributed to data collection. TL, AGS, JS, CH, KN, TC, PA, SV, DM, and ZX contributed to data analysis. TL, AGS, JS, CH, AS, BB, NL, YL, JM, LP, TC, GO, PA, SV, DM, and ZX contributed to drafting and editing the manuscript. All named authors have approved the final manuscript.

## Potential Conflicts of Interests

None.

## References

1. Mahase E. Covid-19: WHO declares pandemic because of “alarming levels” of spread, severity, and inaction. *BMJ* 2020;368:m1036.
2. Ellul MA, Benjamin L, Singh B, et al. Neurological associations of COVID-19. *Lancet neurology* 2020;19(9):767–783.
3. Moriguchi T, Harii N, Goto J, et al. A first case of meningitis/encephalitis associated with SARS-Coronavirus-2. *International journal of infectious diseases* □: *IJID* □: official publication of the International Society for Infectious Diseases 2020;94:55–58.
4. Novi G, Rossi T, Pedemonte E, et al. Acute disseminated encephalomyelitis after SARS-CoV-2 infection. *Neurology: Neuroimmunology & Neuroinflammation* 2020;7(5):e797.
5. Dixon L, Varley J, Gontsarova A, et al. COVID-19-related acute necrotizing encephalopathy with brain stem involvement in a patient with aplastic anemia. *Neurology: Neuroimmunology & Neuroinflammation* 2020;7(5):e789.
6. Bolay H, Gül A, Baykan B. COVID-19 is a Real Headache! *Headache: The Journal of Head and Face Pain* 2020;60(7):1415–1421.
7. Dinkin M, Gao V, Kahan J, et al. COVID-19 presenting with ophthalmoparesis from cranial nerve palsy. *Neurology* 2020;95(5):221–223.
8. Toscano G, Palmerini F, Ravaglia S, et al. Guillain-Barré Syndrome Associated with SARS-CoV-2. *New England Journal of Medicine* 2020;382(26):2574–2576.
9. Dalakas MC. Guillain-Barré syndrome: The first documented COVID-19-triggered autoimmune neurologic disease: More to come with myositis in the offing. *Neurology: Neuroimmunology & Neuroinflammation* 2020;7(5):e781.
10. Paterson RW, Group for the UQSNH for N and NC-19 S, Brown RL, et al. The emerging spectrum of COVID-19 neurology: clinical, radiological and laboratory findings. *Brain* 2020;143(10):awaa240.
11. Lechien JR, Chiesa-Estomba CM, Siaty DRD, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *European Archives of Oto-Rhino-Laryngology* 2020;277(8):2251–2261.

12. Spinato G, Fabbris C, Polesel J, et al. Alterations in Smell or Taste in Mildly Symptomatic Outpatients With SARS-CoV-2 Infection. *JAMA: The Journal of the American Medical Association* 2020;323(20):2089.
13. Rogers JP, Chesney E, Oliver D, 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* 2020;7(7):611–627.
14. Oxley TJ, Mocco J, Majidi S, et al. Large-Vessel Stroke as a Presenting Feature of Covid-19 in the Young. *New England Journal of Medicine* 2020;382(20):e60.
15. Abdel-Mannan O, Eyre M, Löbel U, et al. Neurologic and Radiographic Findings Associated With COVID-19 Infection in Children. *JAMA Neurology* 2020;
16. Mao L, Jin H, Wang M, et al. Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China. *JAMA Neurology* 2020;77(6):683.
17. Varatharaj A, Thomas N, Ellul MA, et al. Neurological and neuropsychiatric complications of COVID-19 in 153 patients: a UK-wide surveillance study. *The Lancet Psychiatry* 2020;7(10):875–882.
18. Benussi A, Pilotto A, Premi E, et al. Clinical characteristics and outcomes of inpatients with neurologic disease and COVID-19 in Brescia, Lombardy, Italy. *Neurology* 2020;95(7):e910–e920.
19. Pleasure SJ, Green AJ, Josephson SA. The Spectrum of Neurologic Disease in the Severe Acute Respiratory Syndrome Coronavirus 2 Pandemic Infection: Neurologists Move to the Frontlines. *JAMA Neurology* 2020;77(6):679.
20. Helms J, Kremer S, Merdji H, et al. Neurologic Features in Severe SARS-CoV-2 Infection. *New England Journal of Medicine* 2020;382(23):2268–2270.
21. Liotta EM, Batra A, Clark JR, et al. Frequent neurologic manifestations and encephalopathy-associated morbidity in Covid-19 patients. *Annals of Clinical and Translational Neurology* 2020;324:603.
22. Brat GA, Weber GM, Gehlenborg N, et al. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. *NPJ digital medicine* 2020;3(1):109.
23. Weber GM, Hong C, Palmer NP, 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* 2020;
24. Klann JG, Weber GM, Estiri H, et al. Validation of a Derived International Patient Severity Algorithm to Support COVID-19 Analytics from Electronic Health Record Data. *medRxiv* 2020;1:2020.10.13.20201855.

25. Wasserman L. All of Statistics: A Concise Course in Statistical Inference. Springer Science & Business Media; 2013.
26. Fisher RA. On the Interpretation of  $\chi^2$  from Contingency Tables, and the Calculation of P. Journal of the Royal Statistical Society 1922;85(1):87–94.
27. Benjamini Y, Drai D, Elmer G, et al. Controlling the false discovery rate in behavior genetics research. Behavioural Brain Research 2001;125(1–2):279–284.
28. Korálnik IJ, Tyler KL. COVID-19: a global threat to the nervous system. Annals of Neurology 2020;
29. Zubair AS, McAlpine LS, Gardin T, et al. Neuropathogenesis and Neurologic Manifestations of the Coronaviruses in the Age of Coronavirus Disease 2019: A Review. JAMA Neurology 2020;77(8):1018–1027.
30. Siegler JE, Cardona P, Arenillas JF, et al. Cerebrovascular events and outcomes in hospitalized patients with COVID-19: The SVIN COVID-19 Multinational Registry. International Journal of Stroke 2020;52:1747493020959216.
31. Zakeri A, Jadhav AP, Sullenger BA, Nimjee SM. Ischemic stroke in COVID-19-positive patients: an overview of SARS-CoV-2 and thrombotic mechanisms for the neurointerventionalist. J Neurointerv Surg 2020;neurintsurg-2020-016794.
32. Bengler M, Williams O, Siddiqui J, Sztrihá L. Intracerebral haemorrhage and COVID-19: Clinical characteristics from a case series. Brain, Behavior, and Immunity 2020;88:940–944.
33. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. The Lancet. Haematology 2020;7(6):e438–e440.
34. Guan W-J, Ni Z-Y, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. New England Journal of Medicine 2020;382(18):1708–1720.
35. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 2020;181(2):271–280.e8.
36. Riquelme C, Acuña MJ, Torrejón J, et al. ACE2 is augmented in dystrophic skeletal muscle and plays a role in decreasing associated fibrosis. PloS one 2014;9(4):e93449.
37. Bagnato S, Boccagni C, Marino G, et al. Critical illness myopathy after COVID-19. International journal of infectious diseases □: IJID □: official publication of the International Society for Infectious Diseases 2020;99:276–278.

38. Zhang Y, Cai T, Yu S, et al. High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). *Nature Protocols* 2019;14(12):3426–3444.

## Figure Legends

**Figure 1. Schematic diagram of the cohort and data generation workflow for each site.**

**Figure 2. Characteristics of the study population across sites 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 site, grouped by country. Node size corresponds to the total number of patients per site. (C) Distribution of self-identified race among patients at sites 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 sites 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.

**Figure 3. Prevalence of neurological phenotypes among all patients.**

(A) Difference in prevalence of each neurological ICD-10 code by site 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 sites. The mean proportion estimates are shown as circles and the 95% confidence intervals are shown as bars.

**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 ( $\hat{\theta}$ ) and expected ( $\cdot$ ) 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.

## Table

**Table 1. Definition of severe COVID-19 based on the 4CE Severity Criteria.**

<b>Severe Illness Category</b>	<b>Clinical Events</b>
Diagnoses	Acute respiratory distress syndrome, ventilator-associated pneumonia
Procedures	Insertion of endotracheal tube; invasive mechanical ventilation
Laboratory results	PaCO <sub>2</sub> , PaO <sub>2</sub>
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

Please see detail of the 4CE severity criteria in a separate manuscript (accepted at Journal of American Medical Informatics):

<https://www.medrxiv.org/content/10.1101/2020.10.13.20201855v1.full>

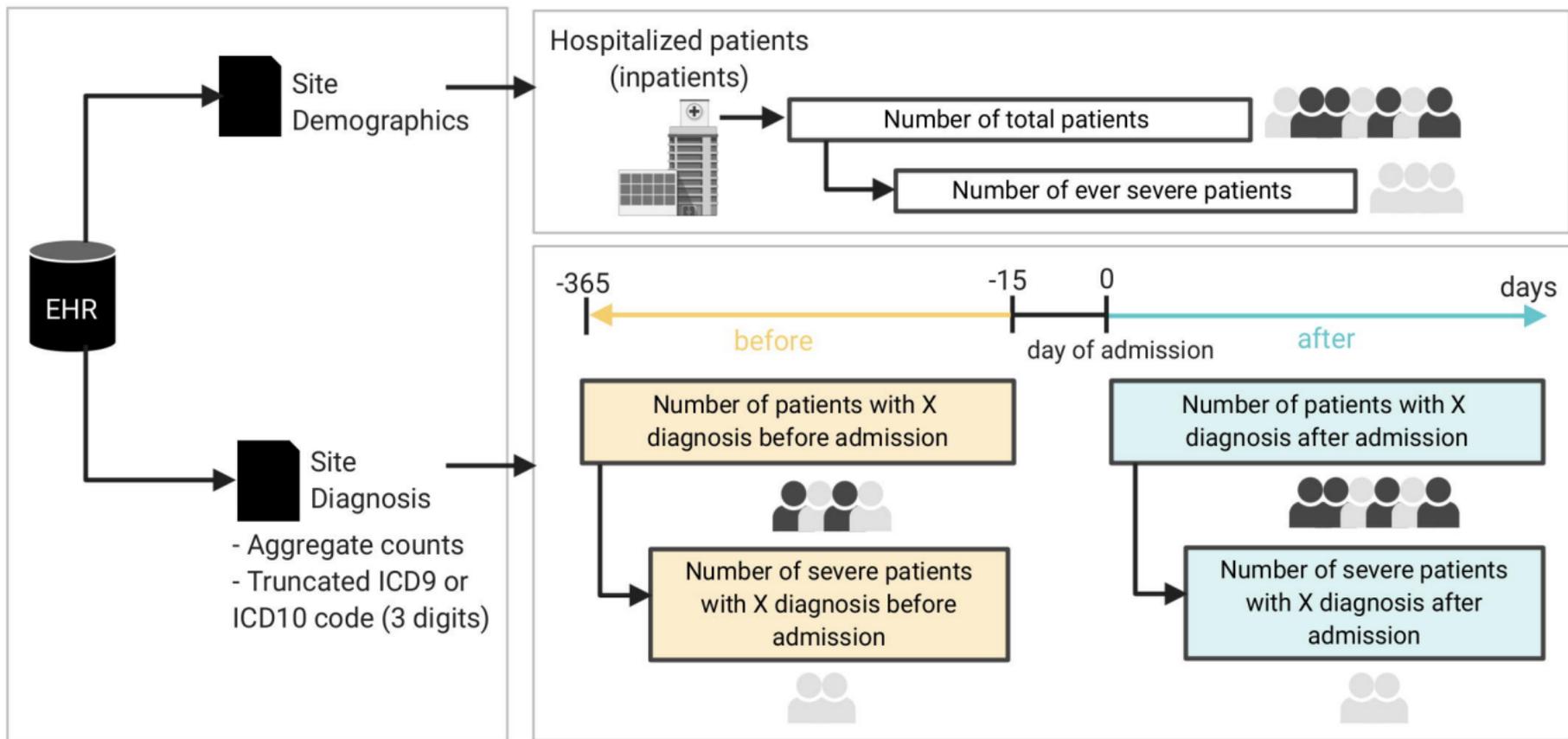
**Table 2. Mapping of neurological disease categories to ICD-10 category codes and their descriptions.**

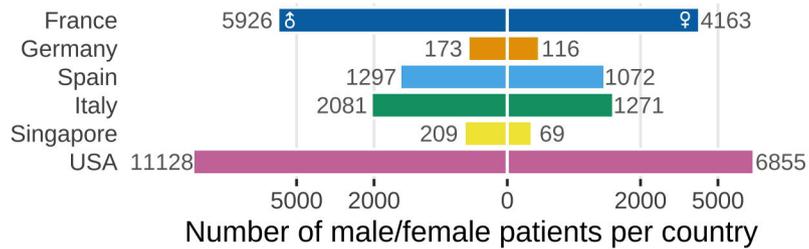
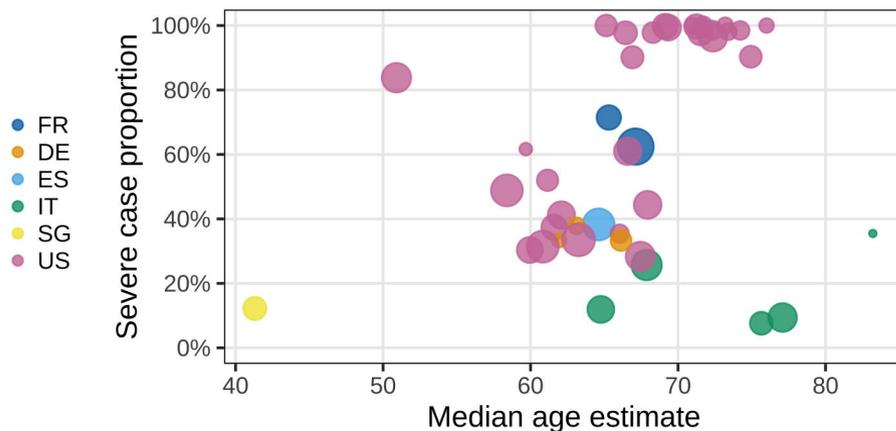
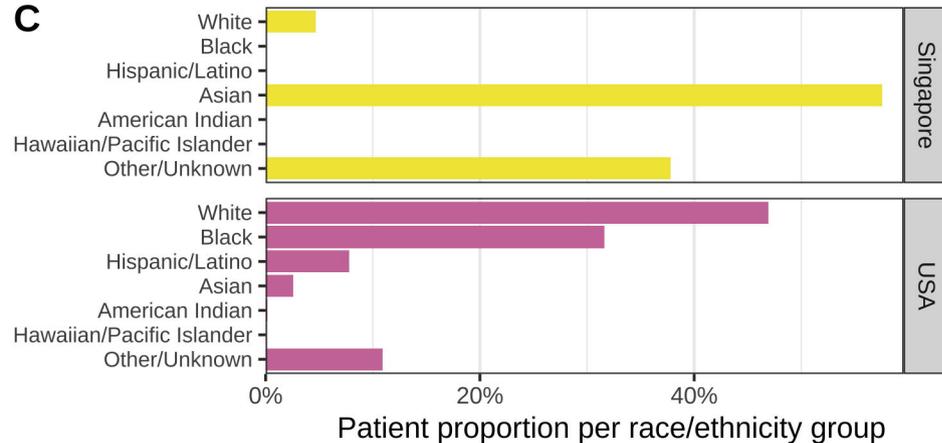
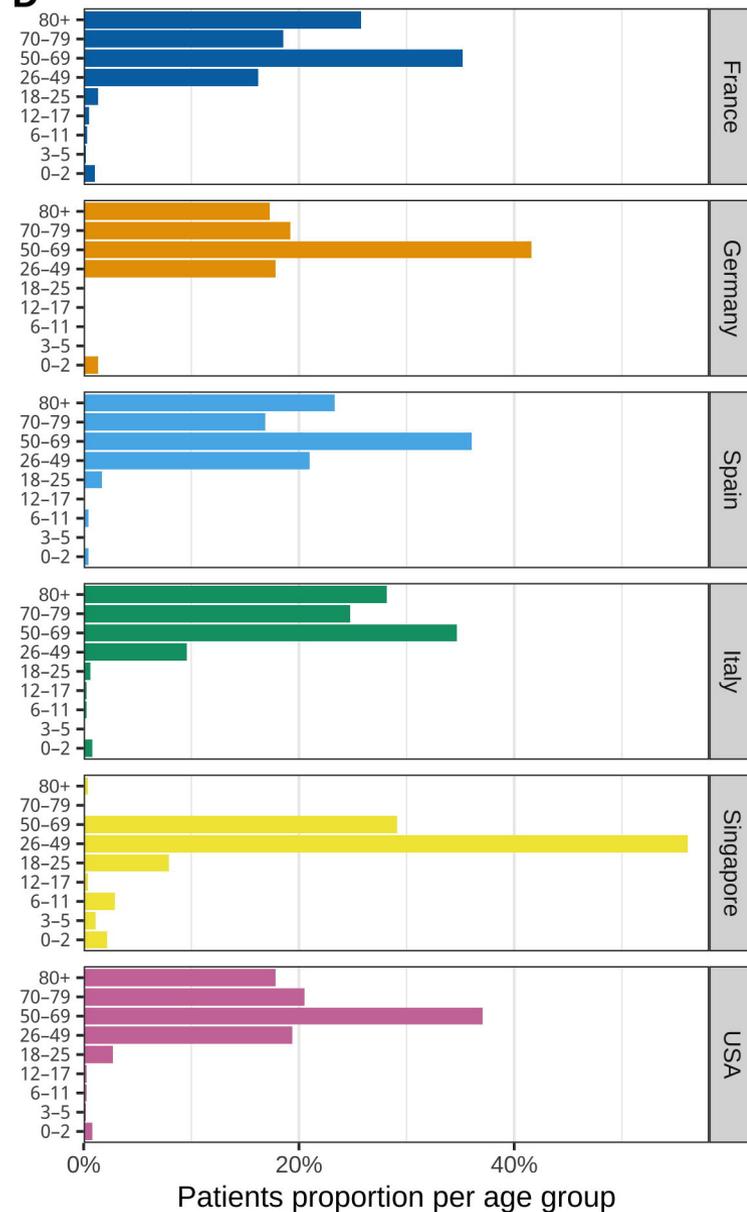
<b>Disease Category</b>	<b>ICD-10 Code</b>	<b>ICD-10 Code Description</b>
Consciousness	R41	Other symptoms and signs involving cognitive functions and awareness
Coordination	R27	Other lack of coordination
Dizziness	R42	Dizziness and giddiness
Headache	G44	Other headache syndromes
Inflammatory	G03	Meningitis 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

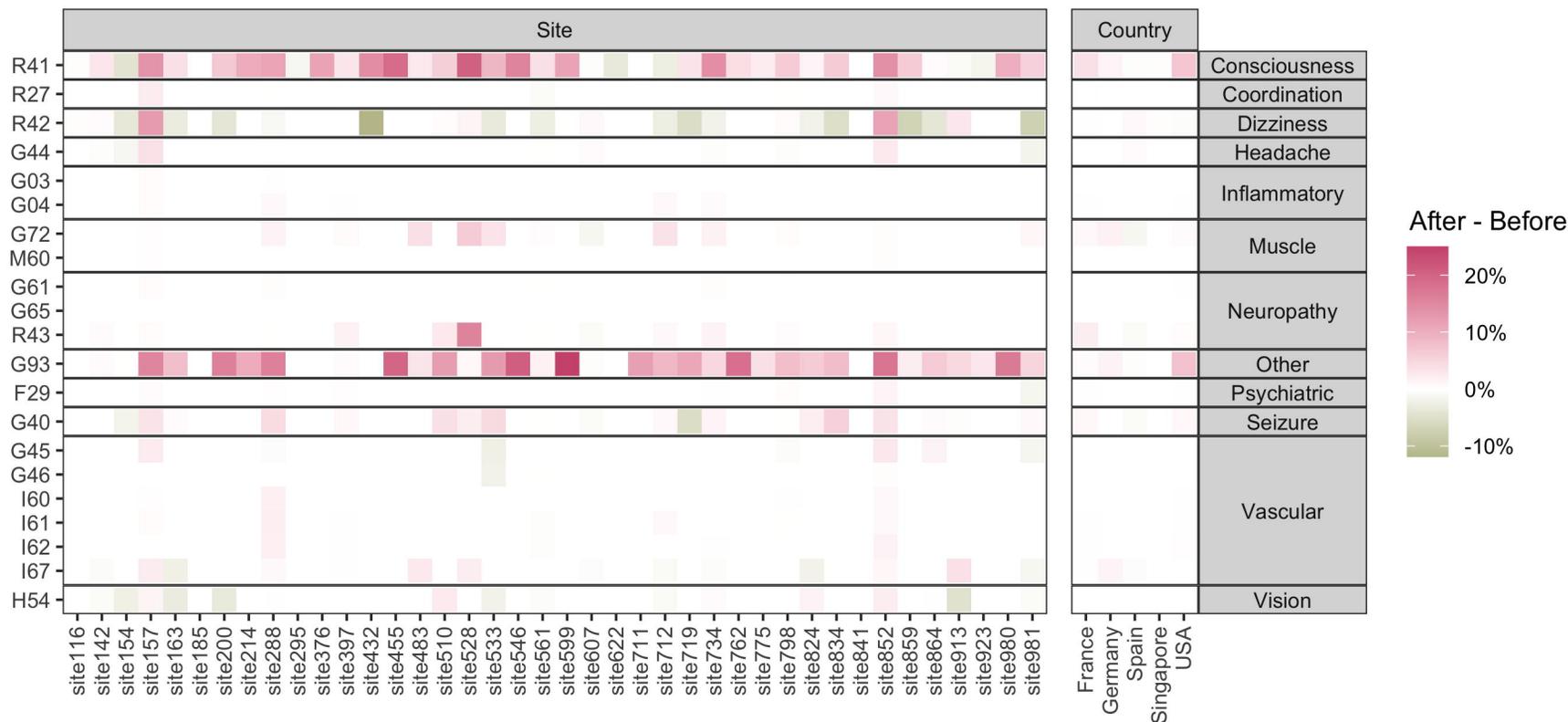
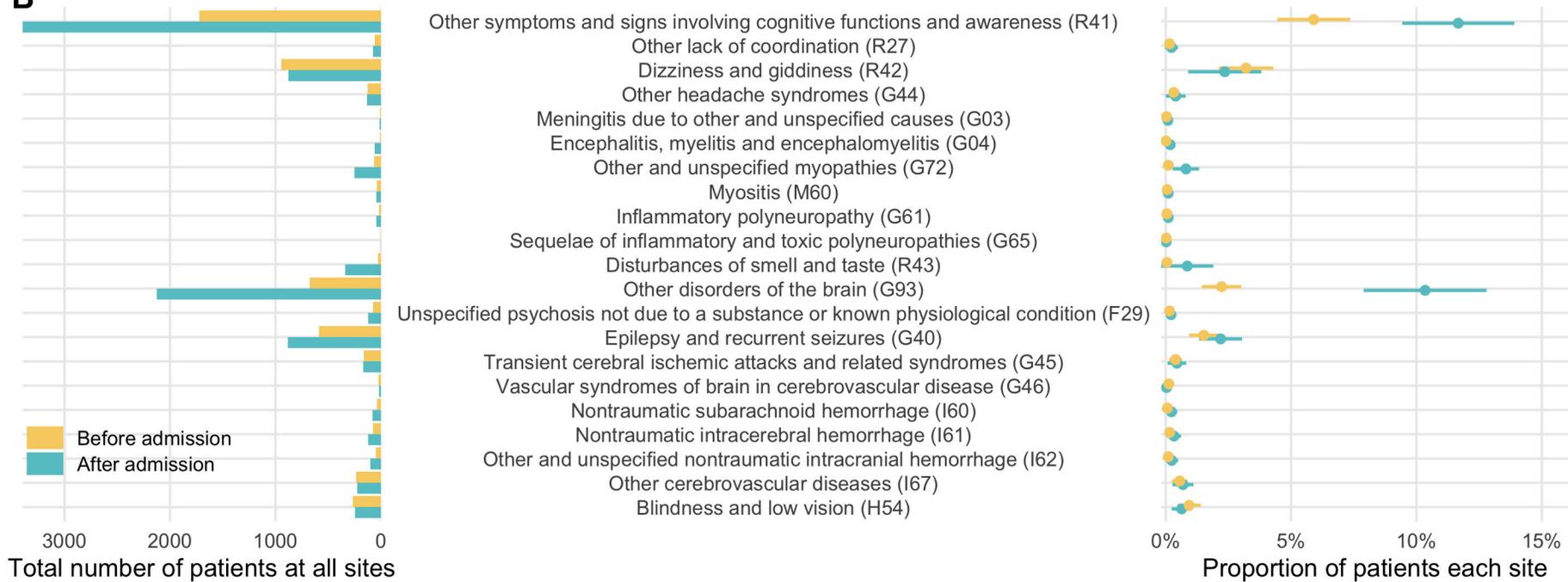
**Table 3.** Statistically significant associations of neurological conditions and severe disease status after admission ( $p_{\text{FDR}} < 0.05$ ).

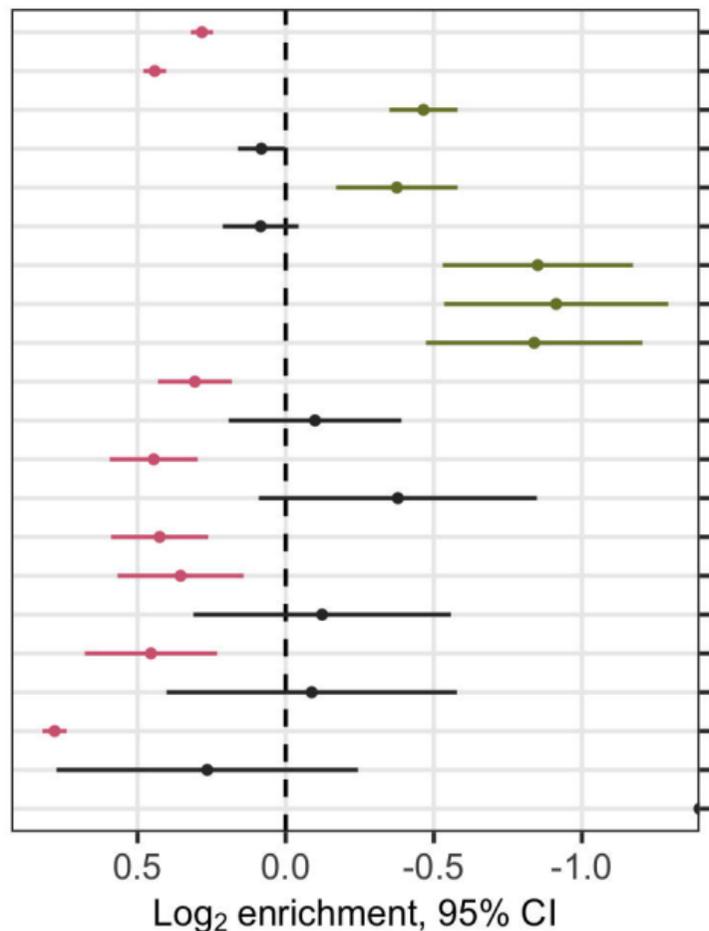
<b>Neurological Condition (ICD-10 Code)</b>	LOE	RRD <sup>a</sup> (%)	RRD <sup>a</sup> 95% CI (%)	$p_{\text{FDR}}$
Blindness and low vision (H54)	-0.38	-23	(-33, -11)	$2.0 \times 10^{-4}$
Dizziness and giddiness (R42)	-0.47	-28	(-33, -22)	$6.3 \times 10^{-19}$
Encephalitis, myelitis and encephalomyelitis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

<sup>a</sup>RRD: Relative risk difference = Observed relative risk - 1.



**A****B****C****D**

**A****B**



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

