Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
**ARTICLE**

**COVID-19 in people with neurofibromatosis 1, neurofibromatosis 2, or schwannomatosis**

Jineta Banerjee\(^1\), Jan M. Friedman\(^2\), Laura J. Klesse\(^3\), Kaleb H. Yohay\(^4\), Justin T. Jordan\(^5\), Scott R. Plotkin\(^6\), Robert J. Allaway\(^1\)*, Jaishri O. Blakeley\(^7\)*

\(^1\)Sage Bionetworks, Seattle, WA; \(^2\)Department of Medical Genetics, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada; \(^3\)Division of Hematology/Oncology, Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX; \(^4\)Departments of Neurology and Pediatrics, NYU Langone Health, New York, NY; \(^5\)Stephen E. and Catherine Pappas Center for Neuro-Oncology, Massachusetts General Hospital, Boston, MA; \(^6\)Stephen E. and Catherine Pappas Center for Neuro-Oncology, Department of Neurology and Cancer Center, Massachusetts General Hospital, Boston, MA; \(^7\)Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD

**ABSTRACT**

**Purpose:** People with pre-existing conditions may be more susceptible to severe COVID-19 when infected by SARS-CoV-2. The relative risk and severity of SARS-CoV-2 infection in people with rare diseases such as neurofibromatosis type 1 (NF1), neurofibromatosis type 2 (NF2), or schwannomatosis (SWN) is unknown.

**Methods:** We investigated the proportions of people with NF1, NF2, or SWN in the National COVID Cohort Collaborative (N3C) electronic health record data set who had a positive test result for SARS-CoV-2 or COVID-19.

**Results:** The cohort sizes in N3C were 2501 (NF1), 665 (NF2), and 762 (SWN). We compared these with N3C cohorts of patients with other rare diseases (98-9844 individuals) and the general non-NF population of 5.6 million. The site- and age-adjusted proportion of people with NF1, NF2, or SWN who had a positive test result for SARS-CoV-2 or COVID-19 (collectively termed positive cases) was not significantly higher than in individuals without NF or other selected rare diseases. There were no severe outcomes reported in the NF2 or SWN cohorts.

**Conclusion:** Having NF1, NF2, or SWN does not appear to increase the risk of being SARS-CoV-2 positive or of being a patient with COVID-19 or of developing severe complications from SARS-CoV-2.

© 2022 Published by Elsevier Inc. on behalf of American College of Medical Genetics and Genomics.
Introduction

Neurofibromatosis type 1 (NF1), neurofibromatosis type 2 (NF2) and schwannomatosis (SWN) (cumulatively termed NF) are autosomal dominant genetic conditions predisposing patients to tumors involving the central and peripheral nervous system. NF1 is much more common (estimated prevalence of 1/3600) than NF2 (1/56,000) or SWN (1/126,000).\(^1,2\)

Given that NF1, NF2, and SWN often cause chronic health impairments, the care community has been concerned about the possibility of increased risk of infection or severe outcomes of COVID-19 in people with one of these genetic conditions. Although these diseases do not generally cause immunosuppression and therefore may not increase susceptibility to SARS-CoV-2 infection, other factors could increase risks associated with COVID-19 in people with NF. For instance, NF1 is associated with several types of malignant tumors (eg, malignant peripheral nerve sheath tumors, juvenile myelomonocytic leukemia, and glioma), nonmalignant tumors, and a range of other manifestations (eg, vasculopathy and cognitive deficits). People with NF1 have a reduced life expectancy attributed predominantly to premature death caused by cancer or vasculopathy.\(^3,4\) Some of these manifestations might increase risks associated with SARS-CoV-2 infection. People with NF1 have also been affected by reduced access to routine care and delayed activity in clinical trials during the COVID-19 pandemic.\(^5,6\) but it is still unknown whether people with NF1, NF2, or SWN are more susceptible to SARS-CoV-2 infection or more likely to have severe symptoms of the disease than other populations.

To address these questions, we explored the electronic health records (EHRs) of people with NF1, NF2, or SWN in the National COVID Cohort Collaborative (N3C) Data Enclave\(^7\) to estimate the proportion of patients with these diagnoses affected by SARS-CoV-2 or COVID-19. The N3C Enclave is a data set and analysis platform that permits researchers to access, query, and analyze COVID-19 related EHR data (including standardized clinical diagnoses, laboratory results, medication records, procedures, and visit records) from 55 participating health care sites and an estimated 6.4 million individuals in the United States (until July 2021)\(^7,8\) to better understand the impact of COVID-19 on specific populations.

This study explored the proportions of patients with a positive test result for SARS-CoV-2 or patients with COVID-19 in those with NF1, NF2, and SWN. Furthermore, we examined the proportions of positive cases with NF1 who experienced high severity of COVID-19 disease based on the retrospective observational data available in N3C.

Materials and Methods

Data access

Data access and analysis were compliant with Sage Bionetworks protocol–granted institutional review board–exempt status by the Western Institutional Review Board - Copernicus Group IR and data access request approved by the N3C Enclave Data Access Committee. All cohorts were generated using custom SQL queries and subsequently analyzed in the N3C Data Enclave. Some of the aggregate data were downloaded after approval by the N3C Data Access Committee and further analyzed in a private, secure cloud computing instance provisioned by Sage Bionetworks. The data analyzed in this study were last updated on July 29, 2021.

Criteria for patients with a positive test result for SARS-CoV-2 or patients with COVID-19

Patients were documented as SARS-CoV-2 positive or as having COVID-19 (together called positive cases) in the N3C Data Enclave if they had a hospital visit after January 1, 2020, and had one or more of the following: (1) a positive result from one or more of a set of predefined SARS-CoV-2 laboratory tests, (2) a strong positive COVID-19 diagnostic code from the International Classification of Diseases, 10th Revision (ICD-10) or Systematized Nomenclature of Medicine (SNOMED) tables described in version 3.3 of the N3C phenotype documentation,\(^9\) or (3) 2 weak positive COVID-19 diagnostic codes from the ICD-10 or SNOMED tables in the phenotype documentation during the same encounter or on the same date before May 1, 2020. Cases in which a patient had both a positive laboratory test result and a positive COVID-19 diagnosis code, priority was given to the criteria that had an earlier date. Cases in which a patient had both positive test results and COVID-19 diagnosis code documented on the same date, priority was given to the positive laboratory test result. Because 1 criterion was selected and documented for each patient to determine their SARS-CoV-2 or COVID-19 status, there was no duplication of individuals if they satisfied >1 criterion.

Each positive case entered in N3C was matched to 2 patients with a negative test result for SARS-CoV-2 (controls) at the same site by age, sex, and race. Control patients met 1 or more of the following criteria: (1) set of predefined SARS-CoV-2 laboratory tests with a non-positive result, (2) did not qualify as a patient with COVID-19, or (3) had at least 10 days between the minimum and maximum encounter date to eliminate patients who were only seen for a COVID test. The N3C cohort and case definition criteria are publicly available as described in version 3.3 of the N3C COVID-19 phenotype documentation (https://github.com/National-COVID-Cohort-Collaborative/Phenotype_Data_Acquisition/wiki/Latest-Phenotype).

Case-control matching at the clinical contributing site

Matching of positive cases and controls was done on-site at the clinical centers before any data were deposited into the N3C Data Enclave. The phenotype and data acquisition workstream of N3C defined the phenotype and provided the
participating sites with programmatic scripts for the matching of patients with positive and negative tests results for SARS-CoV-2. All of the code used to select the cohort and match is available on GitHub (https://github.com/ National-COVID-Cohort-Collaborative/Phenotype_Data_Acquisition/tree/master/PhenotypeScripts). In brief, the code scripts were applied first to select the positive case and then to select potential matched controls at the site depositing records for the positive case. Patients who were qualified to be controls had at least 1 negative COVID test result, never had a positive COVID test result, or did not have U07.1 (COVID) diagnosis code in their record. The scripts iterated through 3 rounds of matching, using progressively looser demographic criteria until either every case had 2 matched controls, or the data set ran out of possible controls. For example, the script attempted to match cases to controls on race, age group, ethnicity, and sex. If 2 matches could not be found using those criteria, then only matching race, age group, and sex was attempted. If 2 matches are still not possible, then matching was done on age and sex. If 2 matches were still not found, then the cases and controls were only matched on sex. At this point, essentially everyone in the case group had 2 matched controls, but it is possible that there may be cases that did not have controls if the site ran out of controls (publication7 and personal communication with Dr. Emily Pfaff, July 07, 2022).

**Patient cohort selection**

Patients included in the N3C data set consist of positive cases and controls from same contributing sites in the ratio 1:2, matched by age, sex (female, male, unknown), race (White, Black or African-American, Native Hawaiian or Pacific Islander, Asian, other, missing/unknown), and ethnicity (Hispanic, non-Hispanic, missing/unknown) (N3C COVID-19 Phenotype version 3.3).7,8

In this study, the data set was stratified by disease, selecting several disease cohorts as comparison groups. An NF1-specific concept set was constructed using NF1-relevant diagnosis codes (Supplemental Table 1) from SNOMED, ICD-9/10, Logical Observation Identifiers Names and Codes (LOINC), and Nebraska Lexicon (N3C Codeset identifier [ID]: 792972142). Any unique person in the N3C data set (identified by their unique N3C person ID) with diagnosis codes belonging to any of the NF1-relevant concepts in the concept set defined by the study team (24 concepts; Supplemental Table 1) was included in the NF1 cohort. This was repeated for NF2 and SWN as well as the other comparison cohorts.

Our comparison cohorts included patients with other rare diseases such as fragile X syndrome (FXS), tuberous sclerosis (TSC), Merkel cell carcinoma (MCC), or acute myeloid leukemia (AML); non-rare diseases such as diabetes mellitus type 1 (DM1) or controlled hypertension (HYP); and the N3C population without NF1 (non-NF1), without NF2 (non-NF2), or without SWN (non-SWN). We selected TSC, FXS, MCC, and AML as comparison groups that are rare diseases and could potentially account for any biases associated with disease groups of small numbers with chronic conditions. TSC, similar to NF1, is an inherited autosomal dominant disease and is closely related to NF1 among neurocutaneous syndromes in both genetic aspects and the presentation of symptoms and diagnosis in childhood.10 Likewise, FXS is a rare X-linked dominant disease that is diagnosed in childhood.11 MCC and AML represent rare cancers12,13 that have molecular similarities to malignancies associated with NF1, NF2, and SWN. We also wanted to compare the rare disease cohorts with other larger disease groups with known association with COVID-19 severity (as positive controls). This led us to select DM1 and HYP groups, which are well-established co-morbidities that showed higher risk of COVID-19 severity. The non-NF1, non-NF2, and non-SWN groups were included to represent the general N3C population.

The concept sets generated for NF1, NF2, and SWN primarily included individuals with confirmed diagnosis of the conditions but also included individuals who presented with conditions associated with these diseases, possibly indicating that they had NF1, NF2, or SWN. This was to ensure that we did not miss an association of the disease and positive cases (and to optimize our cohort selection for sensitivity); a summary of the clinical concepts and their association with the included patients in the NF1, NF2, and SWN cohorts is available in Supplemental Figure 1. All concept sets are available in Supplemental Tables 1 to 9.

Patients with missing data for pre-existing diagnosis, SARS-CoV-2 test, COVID-19 diagnosis, or age were excluded from the analysis. Owing to the data anonymization before contribution to the N3C database, if a patient visited more than 1 of the 55 health care sites contributing to N3C, they would be treated as multiple unique patients (1 per site). This is a known limitation of the data set, but it is unknown whether this scenario occurred in this analysis. In addition, the selected cohorts are not mutually exclusive. A patient with DM1 and HYP were counted in both cohorts. This known limitation of our data set was of low consequence because this study did not aim to look at interaction of diseases but only investigated proportions across any of the selected diseases.

**Considerations and adjustments for the selected cohorts**

The N3C data set contains positive cases and controls in the ratio of 1:2 matched by age, sex, and race. Owing to the inclusion criteria, the analyses in this study cannot accurately estimate the absolute incidence or prevalence of COVID-19 in selected disease cohorts. Additional factors introducing observation bias into the data set include (1) the site contributing data and (2) patients’ age (which contributes to both COVID-19 susceptibility and disease severity as well as the development of symptoms related to NF1, NF2, or SWN and a recorded diagnosis of these conditions).
Site-related adjustment for selected disease cohorts
Contributing sites may introduce observation bias into the contributed data set owing to the nature of patient population, geographic region, or scale of the site. They may also introduce confounding factors into disease severity metrics, eg, there may be variability of criteria in indications for intubation at various health care sites or availability of a given resource or procedure across sites. In addition, data coming from some sites may not include certain variables owing to various reasons such as data not being collected, data not being coded properly, or particular data columns not being submitted to N3C. These variables may be missing for some patients, but the occurrence of such missing variables is not at random and is outside our control. To adjust for these differences and missingness, we included data only from sites that contributed data of patients with NF1, NF2, or SWN. To achieve this, first, the unique data partner IDs (correlating to health care sites) in the NF1, NF2, and SWN cohorts were noted. These data partner IDs were then used to filter all the other cohorts so that only patients contributed by sites that contributed patients with NF1, NF2, or SWN were included.

Age adjustment for selected disease cohorts
The age-related differences between cohorts were adjusted by stratifying the cohorts into 10-year age bins. Each stratum was then weighted using the age-adjusted rate (“aarate”) formula based on US standard (std) population (US Census 2000). Each age-adjusted cohort comprised ages x through y and was calculated using the following formula:

\[
aarate_{x-y} = \sum_{i=x}^{y} \left( \frac{\text{count}_i}{\text{population}_i} \right) \times 100,000 \times \left( \frac{\text{std population}_i}{\sum_{j=x}^{y} \text{std population}_j} \right)
\]

Age-adjusted counts of positive cases, severe outcomes, or invasive ventilation for each cohort, and further details of calculation of age adjustment are available as Supplementary Tables 10 to 12.

Bootstrap analyses
Three main groups of comparisons were assessed: NF1 vs all other study cohorts, DM1 vs all other study cohorts, and NF1 vs all other rare cohorts (TSC, FXS, MCC, AML, NF2, SWN). For the NF1 vs all comparison, a test vector was populated with the age-adjusted proportions from all the separate age strata of the NF1 cohort. A comparison vector was populated with the age-adjusted proportions for all the age strata in all the other cohorts. A Shapiro-Wilk test (base R v3.6.3, shapiro.test function) was used to test the normality of the distributions of the age-adjusted proportions in the different age strata for each cohort (NF1 cohort: \( P = .165 \), Shapiro-Wilk test). The age-adjusted proportions in the test and comparison vectors were compared to estimate the \( P \) value of the real observations (real \( P \) value) (using BSDA v1.2.0, z.test function). Then, the age-adjusted proportions in the test and comparison vectors were resampled 10,000 times to produce 10,000 possible combinations of age-adjusted proportions (using gdata v2.18.0, resample function). For each resampled cohort, a z-test was performed to estimate the distribution of possible \( P \) values generated from the observed proportions. This distribution of \( P \) values generated through bootstrap presents the CIs for the observed real \( P \) value (ie, the null distribution of the \( P \) value). If the real \( P \) value was <.05 but not significantly different from the distribution of various \( P \) values generated in the bootstrap (Wilcoxon rank-sum test), the cohorts in the comparison were considered significantly different with the real \( P \) value unlikely to occur by chance (see Supplemental Methods for more details on the method and rationale).

A similar approach was undertaken for all other comparisons except that a nonparametric Wilcoxon rank-sum test (base R v3.6.3, wilcox.test function) was used to estimate the real \( P \) value for comparisons of severe outcomes and invasive ventilation (owing to non-normal distribution in all cohorts). All real \( P \) values were adjusted to correct for the number of overlapping comparisons for each disease using the Benjamini-Hochberg (BH) method. The distributions of bootstrapped \( P \) values were visualized using R (ggplot2 v3.3.2). Similar analyses were also performed for NF2 and SWN cohorts.

CI calculations
All comparisons were tested using 95% CI as default. In some bootstrap analysis comparisons, the skewed distribution of values did not allow CI calculations at 95% (because the difference between \( \alpha \) achieved from the distribution and \( \alpha_{\text{target}} \) was greater than \( \alpha_{\text{target}}/2 \), where \( \alpha_{\text{target}} = 0.05 \)). In such cases, the highest CI that was able to be calculated is reported (60%). It should be noted that a 60% confidence interval is more likely to reject the null hypothesis than a 95% CI. In this study, CI were calculated and are reported at 95%, and any comparisons with 60% CI have been explicitly noted in the tables.

Results
Demographics of the NF1, NF2, and SWN cohorts are comparable to those of other cohorts in the N3C

From the 6.4 million patients present in N3C Data Enclave (v3.3, July 2021), we selected cohorts of rare (NF1: 2501, NF2: 665, SWN: 762, TSC: 861, AML: 9844, FXS: 98, MCC: 648) and non-rare diseases (non-NF1: 5.6 million,
non-NF2: 5.6 million, non-SWN: 5.6 million, DM1: 66,234, HYP: 1.6 million) using concept sets of EHR diagnosis codes (Table 1; Figure 1A). The FXS cohort was the smallest among the selected cohorts. Other rare disease cohorts were comparable in size but considerably smaller than the non-rare disease cohorts, as expected. The occurrence of NF1, NF2, and SWN cases in the N3C data (NF1: 0.0004 of total N3C patients, NF2: 0.0001 of total N3C patients, SWN: 0.0001 of total N3C patients) was higher than the expected population prevalence of these diseases (NF1: 0.0002,1,25-27 NF2: 0.00002, 1 SWN: 0.000008 2 approximately), indicating that the N3C data set may not represent a random sample of the general population (Supplemental Methods).

The distribution of ages of the selected rare and non-rare disease cohorts was significantly different (median ages provided in Table 1; Kruskal-Wallis rank-sum test, \( P < 2.2 \times 10^{-16} \)). This suggested a need for age adjustment of cohorts before comparison. The NF1, NF2, and SWN cohorts were not significantly different from the non-NF1, non-NF2, or non-SWN cohorts in racial makeup, with a majority of White but a substantial representation from the Black or African-American race (Figure 1B; Supplemental Table 13; Kruskal-Wallis rank-sum test, \( P = .94 \)). The NF1, NF2, SWN, and the general population cohorts did not have significantly different distributions of male and female patients (Figure 1C; Supplemental Table 13; Kruskal-Wallis rank-sum test, \( P = .4159 \)).

### Age-adjusted proportion of SARS-CoV-2–positive cases in NF1, NF2, or SWN is not greater than other selected diseases

To test whether SARS-CoV-2 affected the NF1 population differently than other populations, we compared the age-adjusted proportions of positive cases (patients positive for SARS-CoV-2 and/or patients with COVID-19) in the NF1, NF2, and SWN cohorts individually with that of the non-NF1 population, other rare diseases, and selected non-rare disease cohorts (Figure 2; Supplemental Table 10; Tables 1 and 2). The proportion of positive cases in the NF1 cohort was significantly different from other cohorts (Figure 2A; Tables 1-3; \( z \)-test \( P = .0028 \), BH-adjusted \( P = .008 \)) with a \( P \) value unlikely to occur by chance (see Materials and Methods; bootstrap Wilcoxon \( P = .5 \)). This suggests that the proportion of positive cases in the NF1 N3C cohort was not higher than the non-NF1 N3C cohort. Similarly, the NF2 cohort had a significantly lower proportion of positive cases than the non-NF2 cohort (Figure 2D; Table 2; \( z \)-test \( P = 3.9 \times 10^{-5} \), BH-adjusted \( P = 1.2 \times 10^{-3} \) and falls within bootstrap distribution: Wilcoxon \( P = .5 \)). The age-adjusted proportions of positive cases in the SWN cohort, however, did not differ significantly from the other cohorts (\( z \)-test \( P = .05 \), BH-adjusted \( P = .16 \) and falls within bootstrap distribution: Wilcoxon \( P = 1.0 \); Figure 2G; Table 2).

### Table 1

| Variable Name | NF1 | NF2 | SWN | TSC | AML | FXS | MCC | Non-NF1 | Non-NF2 | Non-SWN | DM1 | HYP |
|---------------|-----|-----|-----|-----|-----|-----|-----|---------|---------|---------|-----|-----|
| Site-adjusted cohort size | 2501 | 665 | 762 | 861 | 9844 | 98 | 648 | 5,577,737 | 5,579,215 | 5,579,146 | 66,234 | 1,664,134 |
| Site-adjusted total positive cases | 352 | 97 | 118 | 149 | 1595 | 26 | 103 | 1,621,928 | 1,622,133 | 1,622,123 | 17,091 | 1,642,134 |
| Median age, y | 28 | 49 | 60 | 23 | 62 | 28 | 75 | 19,823 | 28,503 | 28,498 | 28,498 | 24,949 |
| Age-adjusted counts of positive cases per 100,000 US standard population | 14,496 | 13,782 | 13,738 | 19,346 | 18,135 | 28,105 | 28,498 | 28,503 | 28,498 | 28,498 | 24,022 | 24,949 |
| Approximate percentage of positive cases | 14.5 | 13.7 | 13.8 | 19.3 | 19.8 | 28.1 | 28.5 | 28.5 | 28.5 | 28.5 | 24.0 | 24.9 |
In contrast to NF1 and NF2, the proportion of positive cases observed in a non-rare disease such as DM1 was not significantly different from the rest of the cohorts (z-test $P = .15$, BH-adjusted $P = .15$ and falls within bootstrap distribution: Wilcoxon $P = .8$; Figure 2B, E, and H; Table 2).

We further compared the age-adjusted proportions of positive cases in the NF1 cohort with the other selected rare diseases to test whether the lower proportions of positive cases were unique to NF1. The proportion of positive cases in the NF1 cohort was not significantly different from all the other rare disease cohorts combined (NF2, SWN, non-NF1, non-NF2, and non-SWN cohorts. N3C, National COVID Collaborative Cohort; NF1, neurofibromatosis type 1; NF2, neurofibromatosis type 2; non-NF1, general population without NF1; non-NF2, general population without NF2; non-SWN, general population without SWN; schwannomatosis).

Figure 1  Demographics of selected cohorts in the N3C. A. An example flow diagram showing the various stages of selection of patients to generate the NF1 cohort. The number of patients at each stage is noted. Similar steps were taken during generation of other cohorts in this study. B. Bar plot showing percentage of unique persons who identified as White, Black, or other races in the selected cohorts. C. Bar plot showing percentage of unique persons identifying as male or female in NF1, NF2, SWN, non-NF1, non-NF2, and non-SWN cohorts. N3C, National COVID Collaborative Cohort; NF1, neurofibromatosis type 1; NF2, neurofibromatosis type 2; non-NF1, general population without NF1; non-NF2, general population without NF2; non-SWN, general population without SWN; schwannomatosis.
Figure 2  Comparison of age-adjusted proportions of positive cases in selected cohorts. Selected cohorts included NF1, tuberous sclerosis, acute myeloid leukemia, fragile X syndrome, Merkel cell carcinoma, non-NF1/non-NF2/non-SWN (general population), DM1, and controlled hypertension. A-C. Results of the bootstrap analysis for \( P \) value of comparisons, NF1 vs all, DM1 vs all, NF1 vs rare disease cohorts. D-F. Results of the bootstrap analysis for comparisons, NF2 vs all, DM1 vs all, NF2 vs rare disease cohorts. G-I. Results of the bootstrap analysis for comparisons of \( P \) values of SWN vs all, DM1 vs all, SWN vs rare disease cohorts. The red dashed line represents the \( P \) value obtained from the real observations. The specific and adjusted value for each comparison is noted in the plot inset. The gray bars show a histogram of all possible \( P \) values obtained through 10,000 iterations of bootstrap analysis. adj, adjusted; DM1, diabetes mellitus type 1; NF1, neurofibromatosis type 1; NF2, neurofibromatosis type 2; SWN, schwannomatosis.
Table 2  Table showing age-adjusted counts and $P$ values of all comparisons of positive cases in selected cohorts

| Comparison     | Age-Adjusted Counts of Positive Cases in Target Disease (per 100,000) | $z$-Test $P$ Value | Estimate (Difference in Mean Between Groups) | 95% CI (Lower) | 95% CI (Upper) | Adjusted $z$-Test $P$ Value | Wilcoxon $P$ Value (Bootstrap) | Estimate (Difference in Location Parameter Between Groups) | 60% CI (Lower) of Estimate (Bootstrap) | 60% CI (Upper) of Estimate (Bootstrap) |
|----------------|---------------------------------------------------------------|--------------------|---------------------------------------------|---------------|---------------|-----------------------------|-------------------------------|--------------------------------------------------|-------------------------------------|-------------------------------------|
| NF1 vs all     | 14,496                                                        | .002               | −832.47                                     | −1380.16      | −284.80        | .008                        | .5                             | −0.26                              | −0.97                               | −0.19                               |
| NF2 vs all     | 13,782                                                        | .00003             | −931.63                                     | −1375.85      | −687.41        | .00012                      | .5                             | −0.34                              | −0.60                               | −0.11                               |
| SWN vs all     | 13,738                                                        | .05                | −937.65                                     | −1894.27      | 18.95          | .16                         | 1                              | 0.002                              | −0.74                               | 0.005                               |
| DM1 vs all     | 24,022                                                        | .15                | 490.63                                      | −184.29       | 1165.70        | .15                         | .8                             | 0.09                               | −0.23                               | 0.15                                |
| NF1 vs rare    | 14,496                                                        | .08                | −540.66                                     | −1153.10      | 71.76          | .12                         | .4                             | −0.10                              | −0.79                               | 0.05                                |
| NF2 vs rare    | 13,782                                                        | .015               | −644.84                                     | −1167.71      | −121.98        | .02                         | .4                             | −0.06                              | −0.54                               | 0.01                                |
| SWN vs rare    | 13,738                                                        | .19                | −651.17                                     | −1641.66      | 339.31         | .19                         | 1                              | 0.02                               | −0.71                               | 0.19                                |

In the bootstrap analysis, the $z$-test $P$ value was compared with a distribution of bootstrapped $P$ values using the nonparametric Wilcoxon rank-sum test. In this test, the null hypothesis is that the 2 distributions differ by a location shift of $\mu$ (magnitude of the location shift), and the alternative hypothesis is that they differ by a location shift other than $\mu$. The estimate of this nonparametric test is equal to the difference in $\mu$, which in this case has negative values because of the direction of the location shift. The CIs reflect the range of $\mu$ and have negative values. The skew in the distribution of bootstrapped $P$ values did not allow CI calculations at 95% (because the difference between $\alpha$ achieved from the distribution and $\alpha_{\text{target}}$ was greater than $\alpha_{\text{target}}/2$, where $\alpha_{\text{target}} = 0.05$) but enabled 60% CI estimation. A 60% CI is more likely to reject the null hypothesis than a 95% CI.

$DM1$, diabetes mellitus type 1; $NF1$, neurofibromatosis type 1; $NF2$, neurofibromatosis type 2; $SWN$, schwannomatosis.

Table 3  Table showing age-adjusted counts, CI, and $P$ values of all comparisons of severe outcomes and invasive ventilation recorded in selected cohorts

| Comparison     | Age-Adjusted Counts of Severe Outcomes in Target Disease (per 100,000) | Wilcoxon $P$ Value | Estimate (Difference in Location Parameter Between Groups) | 95% CI (Lower) of Estimate | 95% CI (Upper) of Estimate | Adjusted Wilcoxon $P$ Value | Wilcoxon $P$ Value (Bootstrap) | Estimate (Difference in Location Parameter Between Groups) (Bootstrap) | CI (Lower) of Estimate (Bootstrap) | CI (Upper) of Estimate (Bootstrap) |
|----------------|-------------------------------------------------------------------------|--------------------|----------------------------------------------------------|----------------------------|---------------------------|-----------------------------|-------------------------------|--------------------------------------------------------------------------------|-----------------------------------|-----------------------------------|
| Severe outcomes|                                                                           |                    |                                                          |                            |                           |                             |                               |                                                                                |                                  |                                   |
| NF1 vs all     | 887                                                                      | .56                | 0.0000016                                                | −0.00006                   | 204.70                    | .56                         | .8                            | −0.18                           | −0.43a                              | −0.04a                             |
| DM1 vs all     | 534                                                                      | .003               | 42.10                                                    | 28.05                      | 64.43                     | .009                        | 1                             | −0.07                           | −0.36a                              | 0.002a                             |
| NF1 vs rare    | 887                                                                      | .04                | 0.000002                                                 | −0.000019                  | 204.70                    | .06                         | .28                           | −0.04                           | −0.34                               | −0.034                             |
| Invasive ventilation|                                                                 |                    |                                                          |                            |                           |                             |                               |                                                                                |                                  |                                   |
| NF1 vs all     | 521                                                                      | .917               | −0.0000054                                               | −0.00008                   | 0.00001                  | .92                         | 1                             | 0.10                            | −0.08a                              | 0.28a                              |
| DM1 vs all     | 303                                                                      | .0002              | 32.25                                                    | 28.05                      | 47.28                     | .0006                       | .4                             | −0.03                           | −0.18a                              | −0.002a                             |
| NF1 vs rare    | 521                                                                      | .06                | 0.0000038                                                | −0.000040                  | 0.00                      | .07                         | .28                           | −0.28                           | −0.93                               | −0.1                                |

In some cases, the skew in the distribution of bootstrapped $P$ values did not allow CI calculations at 95% CI (because the difference between $\alpha$ achieved from the distribution and $\alpha_{\text{target}}$ was greater than $\alpha_{\text{target}}/2$, where $\alpha_{\text{target}} = 0.05$). The confidence limits without superscript $^a$ denote 95% CI. A 60% CI is more likely to reject the null hypothesis than a 95% CI.

$DM1$, diabetes mellitus type 1; $NF1$, neurofibromatosis type 1.

$^a$CI calculations at the 60% level are reported for these comparisons.
rare disease cohorts (z-test \( P = .015 \), BH-adjusted \( P = .02 \) and falls within bootstrap distribution: Wilcoxon \( P = .4 \); Figure 2F; Table 2).

Interpreting these results conservatively, the proportions of positive cases in the N3C NF1, NF2, and SWN populations were no greater than expected for rare or non-rare diseases examined in this study.

**Age-adjusted proportion of severe outcomes in NF1 was not greater than that of other diseases**

Although the positive cases did not seem to be more frequent in people with NF1, NF2, and SWN versus others, it is possible that the severity of COVID-19 in positive cases with NF1, NF2, or SWN is different from the others. In the N3C cohort, there were no patients with NF2, SWN, FXS, or TSC who had reported severe outcomes; thus, we were unable to evaluate the prevalence of severe outcomes in NF2 or SWN cohorts. We evaluated the severity of COVID-19 manifestations in the NF1 cohort and compared that with other selected cohorts. N3C has made extensive efforts to capture the severity of disease by incorporating information from EHRs such as hospitalization, invasive ventilation, extracorporeal membrane oxygenation, hospice, and death.8

We examined patient severity scores built on these parameters in our selected cohorts to estimate the severity of COVID-19.

We first identified the patients in the previously examined disease cohorts with highest documented severity,28 (N3C severity type severe, ie, World Health Organization severity 7-9 and mortality/hospice, ie, World Health Organization severity 10). We then compared the age-adjusted proportions of patients with these severity types (henceforth referred to as severe outcomes) among positive cases in each cohort (Supplemental Figure 2A-C; Supplemental Table 11). We found that the proportion of severe outcomes in the NF1 cohort was not significantly different when compared with all other cohorts (Wilcoxon test, observed \( P = .56 \), BH-adjusted \( P = .56 \) and falls within bootstrap distribution: \( P = .8 \); Supplemental Figure 2A; Table 3). In contrast, we found that the DM1 cohort had higher proportions of patients with severe outcomes than the other selected cohorts (Wilcoxon rank-sum test, observed \( P = .003 \), BH-adjusted \( P = .009 \) and falls within bootstrap distribution: \( P = 1.0 \); Supplemental Figure 2B; Table 3). This finding is consistent with the now established association between diabetes and severity of COVID-19 outcomes.29,30 The proportion of patients with severe outcomes in the NF1 cohort was not significantly higher than that in the other rare disease cohorts examined, suggesting no clear relationship between NF1 and severe outcome from COVID-19 infection/disease (Wilcoxon test, observed \( P = .04 \), BH-adjusted \( P = .06 \) and falls within bootstrap distribution \( P = .28 \); Supplemental Figure 2C; Table 3).

We also examined the proportions of positive cases in the NF1 cohort who received invasive ventilation (Supplemental Figure 2D-F; Table 3; Supplemental Table 12). The proportions requiring invasive ventilation among the positive cases in the NF1 cohort were not significantly different from that in the other cohorts (Wilcoxon test, observed \( P = .91 \), BH-adjusted \( P = .92 \), falls within bootstrap distribution: \( P = 1.0 \); Supplemental Figure 2D and F; Table 3). In contrast, the DM1 cohort appears to have more invasive ventilation (Wilcoxon test, observed \( P = .0002 \), BH-adjusted \( P = .0006 \); Supplemental Figure 2E; Table 3). The median length of hospital stays for the patients with NF1 who had severe outcomes was not substantially different from that for the patients in other cohorts (NF1: 10 days, AML: 9 days, MCC: 23 days, non-NF1: 11 days, DM1: 13 days, HYP: 11 days, TSC: not determined, FXS: not determined, NF2: not determined, SWN: not determined; Wilcoxon test \( P = .55 \); Supplemental Table 14).

Thus, our findings suggest that the proportion of positive cases in the NF1 cohort experiencing severe outcomes was not significantly greater than the non-rare or rare disease cohorts.

**Discussion**

This study examined EHR data in the N3C Data Enclave to determine the burden of SARS-CoV-2 in people with NF1, NF2, and SWN. Our findings suggest that the proportion of cases of SARS-CoV-2 infection or COVID-19 and severe outcomes among patients with NF1, NF2, and SWN in the N3C Data Enclave was not higher than other selected rare and non-rare diseases when adjusted for age and site.

The N3C is the largest centralized and harmonized EHR repository of a representative COVID-19 cohort in the United States,8 which is well suited for studying COVID-19–related outcomes. Although an extensive collection of EHR data, N3C data set has various limitations in determining SARS-CoV-2 and COVID-19–related risks for the global population. For example, N3C shows a greater NF1 prevalence than the population prevalence estimates, indicating that this data set may not represent the general population owing to its specific data acquisition protocols. As with any multisite combination of EHR data, there may also be site-related differences in clinical measures due to the variations in clinical practice and medical record documentation. Any biases affecting analyses of care patterns or outcomes due to geographic, regional, cultural, or other differences between institutions remained unassessed because of anonymized coding data from contributing institutions in this de-identified data set. Furthermore, clinical coding of patients with NF1, NF2, SWN, or other rare diseases may be incomplete in EHRs, ie, the N3C may miss patients with disease without appropriately recorded diagnostic codes. Conversely, patients with rare diseases who are control cases may be more likely to visit the health care system than the people who are control cases but belong to other disease cohorts, thus affecting the proportion estimates.
in this study. Finally, the SARS-CoV-2 testing rate, COVID-19 diagnosis and treatment, access to clinical care in different sites, and disease populations may vary.

In addition, there are various data-specific considerations that could also introduce bias. Owing to N3C’s phenotype acquisition design, the patients included in the N3C were matched by demographics and not by disease type. This matching strategy suggests that the N3C may not provide an accurate proportion of the SARS-CoV-2-negative/ non–COVID-19 population for different disease cohorts, limiting the comparisons to those across different disease cohorts within the N3C. In addition, because the matching of 2 controls to each positive case was done at the time of data deposition at each participating site and occurred in 3 rounds using progressively looser requirements for the match (see case-control matching in Materials and Methods), there may be subtle differences in the cohorts despite the best efforts of the phenotype acquisition team in the N3C. It is also important to note that the definition of patients positive for SARS-CoV-2 or COVID-19 in this study is subject to limitations such as potential false positives, false negatives, and untested asymptomatic individuals. In the presence of at-home testing, it is also likely that people identified as controls may have had COVID-19 but tested outside of their health system or at home. Owing to this caveat, which is not addressable within the construct of the N3C study, it is necessary to assume that the controls are patients who have not had documented instances of COVID-19. This data set also lacks records of individuals who did not have a clinical encounter at N3C contributing sites because of being suspected positive or asymptomatic. Furthermore, this study only focused on the acute data related to SARS-CoV-2 infections and associated critical care use. Future analyses evaluating additional patient covariates known to affect SARS-CoV-2 outcomes, such as pregnancy, social determinants of health (including socioeconomic status, insurance status, and access to health care), or long COVID data, may help us refine the results of this study. Finally, sample sizes for certain diseases (eg, FXS) make interpretation of the results for these diseases challenging. Additional data for the cohorts may also allow higher CI estimates for comparisons in the study, whereas these data only enabled 60% CI calculations.

Despite these limitations, we recapitulated well-established associations between COVID-19 and DM1 (one of our control cohorts), suggesting that our methods can identify underlying patterns of SARS-CoV-2 risk and severity in a common disease. Similar to previous observations, our findings suggest that although the proportion of patients with DM1 found to be positive cases is not greater than the general population (Figure 2C), the proportion of positive cases in the DM1 cohort experiencing severe outcomes was higher than the rest of the comparison cohorts (Supplemental Figure 2B and E; Tables 2 and 3). These observations in our analysis are reassuring and indicate that our analytical methods may be tolerant of the inherent biases and limitations of the N3C data set and EHR data while identifying robust patterns for selected cohorts.

Very few studies have investigated the effect of COVID-19 on rare diseases because of the challenges in collecting information regarding rare disease cohorts. Recently, a prospective multicenter questionnaire study (4 centers) was published with 48 patients with lysosomal storage disorders presenting a descriptive assessment of the effect of the pandemic restrictions on the disease population and their treatment adherence. Similar to our study and other rare disease studies, the authors of this study also struggled with the limited size of the cohort. In addition, they highlighted that (1) the study assessed patients at a time when the health care system had not yet faced massive influx of patients with COVID-19 and (2) there was absence of a control cohort, such as healthy individuals or patients harboring other chronic conditions. The authors suggested that evaluating larger cohorts of patients and comparing them with relevant control cohorts was essential. As a result, this study focused mainly on descriptive analysis of the data from their cohort.

Our study is also a multicenter study (59 clinical sites) that used observational retrospective data from different rare disease cohorts to quantitatively assess the occurrence of COVID-19. We were fortunate to overcome many of the above limitations owing to the centralized efforts of the N3C and present a quantitative assessment of the occurrence of COVID-19 in patients with rare diseases. First, although our cohorts were still very small (requiring complex statistical analyses), they were still significantly larger than the cohorts examined earlier for rare diseases. Second, this study assessed patient data obtained between 2018 and 2021, thus capturing the peak of the COVID-19–related patient influx and the early stages of the pandemic. Our study also struggled with evaluating populations who never had COVID-19 or SARS-CoV-2 (given the lack of data from at-home testing or testing outside the clinical sites included in the study) but was able to contrast the cohorts of patients with NF1, NF2, and SWN with cohorts of other rare diseases such as FXS, AML, MCC, as well as non-rare diseases such as DM1 and HYP.

In the future, additional studies (eg, case-control health surveys, mobile health studies) may allow more accurate determination of the prevalence of COVID-19 and its impact on health in people with NF1, NF2, and SWN. Future studies should also evaluate whether people with rare diseases exhibit cautious behavior or stronger adherence to social distancing protocols contributing to lower SARS-CoV-2 infection rates. This study leverages a new and unique data set and overcomes various statistical challenges to assess COVID-19 burden and severity in a rare disease population. We anticipate that the strategies used in this study can be easily extended to examine other rare diseases of interest using the N3C data set, thus serving as a roadmap for future work. As noted in the results, the number of patients with NF1 permitted more granular analyses than that of those with NF2 and SWN, suggesting that cohort sizes of >1500 individuals from the N3C database are perhaps more feasible to study than those of <1000. As more data are added to the
N3C data set and more concept sets of rare diseases are generated and validated by clinical experts, we anticipate that the role of COVID-19 in rare diseases will become easier to evaluate using the N3C. Using these methods, we discovered that people with NF1, NF2, and SWN do not seem to be at a greater risk of becoming positive cases or developing severe complications of COVID-19 than those with other rare or non-rare diseases. These findings suggest that although no elevated risk was noted as per the composition of N3C patient population in July 2021, it is important for people with NF to follow COVID-19-related public health measures, vaccination guidelines, and recommendations from NF specialists.

**Data Availability**

Data are available in the N3C Data Enclave (https://covid.cd2h.org/enclave). All R code used in the analyses in this study are available on GitHub (https://github.com/jaybee84/NF-COVID-response).

**Individual acknowledgments for core contributors**

The authors gratefully acknowledge the contributions from the following N3C core teams (asterisks indicate leads):

- Principal investigators: Melissa A. Haendel*, Christopher G. Chute*, Kenneth R. Gersing, Anita Walden
- Workstream, subgroup, and administrative leaders: Melissa A. Haendel*, Tellen D. Bennett, Christopher G. Chute, David A. Eichmann, Justin Guinney, Warren A. Kibbe, Hongfang Liu, Philip R.O. Payne, Emily R. Pfaff, Peter N. Robinson, Joel H. Saltz, Heidi Spratt, Justin Starren, Christine Suver, Adam B. Wilcox, Andrew E. Williams, Chunlei Wu
- Key liaisons at data partner sites
- Regulatory staff at data partner sites
- Individuals at the sites who are responsible for creating the data sets and submitting data to the N3C
- Data ingest and harmonization team: Christopher G. Chute*, Emily R. Pfaff*, Davera Gabriel, Stephanie S. Hong, Kristin Kostka, Harold P. Lehmann, Richard A. Moffitt, Michele Morris, Matvey B. Palchuk, Xiaohan Tanner Zhang, Richard L. Zhu
- Phenotype team (individuals who create the scripts that the sites use to submit their data, based on the COVID and long COVID definitions): Emily R. Pfaff*, Benjamin Amor, Mark M. Bissell, Marshall Clark, Andrew T. Girvin, Stephanie S. Hong, Kristin Kostka, Adam M. Lee, Robert T. Miller, Michele Morris, Matvey B. Palchuk, Kellie M. Walters
- Project management and operations team: Anita Walden*, Yooree Chae, Connor Cook, Alexandra Dest, Racquel R. Dietz, Thomas Dillon, Patricia A. Francis, Rafael Fuentes, Alexis Graves, Julie A. McMurry, Andrew J. Neumann, Shawn T. O’Neil, Usman Sheikh, Andréa M. Volz, Elizabeth Zampino
- Partners from the NIH and other federal agencies: Christopher P. Austin*, Kenneth R. Gersing*, Samuel Bozzette, Mariam Deacy, Nicole Garbarini, Michael G. Kurilla, Sam G. Michael, Joni L. Rutter, Meredith Temple-O’Connor
- Analytics team (individuals who build the Enclave infrastructure, help create codesets and variables, and help domain teams and project teams with their data sets): Benjamin Amor*, Mark M. Bissell, Katie Rebecca Bradwell, Andrew T. Girvin, Amini Manna, Nabeel Qureshi
- Publication committee management team: Mary Morrison Saltz*, Christine Suver*, Christopher G. Chute, Melissa A. Haendel, Julie A. McMurry, Andréa M. Volz, Anita Walden
- Publication committee review team: Carolyn Bramante, Jeremy Richard Harper, Wenndy Hernandez, Farrukh M. Koraishy, Federico Mariona, Amit Saha, Satyantarayana Vedula

**Data partners with released data**

Stony Brook University (U24TR002306); Oklahoma Clinical and Translational Science Institute, The University of Oklahoma Health Sciences Center (U54GM104938); West Virginia Clinical and Translational Science Institute, West Virginia University (U54GM104942); Mississippi Center for Clinical and Translational Research, University of Mississippi Medical Center (U54GM115428); Great Plains IDeA-Clinical and Translational Research, University of Nebraska Medical Center (U54GM115458); Northern New England Clinical & Translational Research (NNE-CTR) Network, Maine Medical Center (U54GM115516); Wake Forest Clinical and Translational Science Institute, Wake Forest University Health Sciences (UL1TR001420); Northwestern University Clinical and Translational Science Institute, Northwestern University at Chicago (UL1TR001422); Center for Clinical and Translational Science & Training, University of Cincinnati (UL1TR001425); The Institute for Translational Sciences, The University of Texas Medical Branch at Galveston (UL1TR001439); South Carolina Clinical & Translational Research Institute, Medical University of South Carolina (UL1TR001450); The UMass Center for Clinical and Translational Science, University of Massachusetts Medical School Worcester (UL1TR001453); The Southern California Clinical and Translational Science Institute, University of Southern California (UL1TR001855); Irving Institute for Clinical and Translational Research, Columbia University Irving Medical Center (UL1TR001873); Clinical and Translational Science Institute at Children’s National, George Washington Children’s Research Institute (UL1TR001876); UC Center for Clinical and Translational Science, University of Kentucky (UL1TR001998); UR Clinical & Translational Science Institute, University of Rochester (UL1TR002001); UIC Center for...
Clinical and Translational Science, University of Illinois at Chicago (UL1TR002003); Penn State Clinical and Translational Science Institute, Penn State Health Milton S. Hershey Medical Center (UL1TR002014); Michigan Institute for Clinical & Health Research, The University of Michigan at Ann Arbor (UL1TR002240); Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center (UL1TR002243); Institute of Translational Health Sciences, University of Washington (UL1TR002319); Institute of Clinical and Translational Sciences, Washington University in St. Louis (UL1TR002345); Oregon Clinical and Translational Research Institute, Oregon Health & Science University (UL1TR002369); UW Institute for Clinical and Translational Research, University of Wisconsin-Madison (UL1TR002373); The Institute for Translational Medicine (ITM), Rush University Medical Center (UL1TR002389); The Institute for Translational Medicine (ITM), The University of Chicago (UL1TR002389); North Carolina Translational and Clinical Science Institute, University of North Carolina at Chapel Hill (UL1TR002489); Clinical and Translational Science Institute, University of Minnesota (UL1TR002494); Colorado Clinical and Translational Sciences Institute, Children’s Hospital Colorado (UL1TR002535); Institute for Clinical and Translational Science, The University of Iowa (UL1TR002537); Uhealth Center for Clinical and Translational Science, The University of Utah (UL1TR002538); Tufts Clinical and Translational Science Institute, Tufts Medical Center (UL1TR002544); Duke Clinical and Translational Science Institute, Duke University (UL1TR002553); C. Kenneth and Dianne Wright Center for Clinical and Translational Research, Virginia Commonwealth University (UL1TR002649); Center for Clinical and Translational Science, The Ohio State University (UL1TR002733); University of Miami Clinical and Translational Science Institute, The University of Miami Leonard M. Miller School of Medicine (UL1TR002736); integrated Translational Health Research Institute of Virginia, University of Virginia (UL1TR003015); integrated Translational Health Research Institute of Virginia, Carilion Clinic (UL1TR003015); Center for Clinical and Translational Science, University of Alabama at Birmingham (UL1TR003096); Johns Hopkins Institute for Clinical and Translational Research, Johns Hopkins University (UL1TR003098); UAMS Translational Research Institute, University of Arkansas for Medical Sciences (UL1TR003107); Delaware CTR ACCEL Program, Nemours (U54GM104941); Louisiana Clinical & Translational Science Center, University Medical Center New Orleans (U54GM104940); Colorado Clinical and Translational Sciences Institute, University of Colorado Denver, Anschutz Medical Campus (UL1TR002535); Mayo Clinic Center for Clinical and Translational Science (CCaTS), Mayo Clinic Rochester (UL1TR002377); Center for Clinical and Translational Science, Tulane University (UL1TR003096); The Institute for Translational Medicine (ITM), Loyola University Medical Center (UL1TR002389); The Institute for Translational Medicine (ITM), Advocate Health Care Network (UL1TR002389); Bill and Melinda Gates Foundation grant to Sage Bionetworks, OCHIN (INV-018455).

Additional data partners who have signed a data transfer agreement and whose data release is pending

Center for Clinical and Translational Science, The Rockefeller University (UL1TR001866); Scripps Research Translational Institute, The Scripps Research Institute (UL1TR002550); Institute for Integration of Medicine and Science, University of Texas Health Science Center at San Antonio (UL1TR002645); Center for Clinical and Translational Sciences, The University of Texas Health Science Center at Houston (UL1TR003167); The Institute for Translational Medicine (ITM), NorthShore University HealthSystem (UL1TR002389); Yale Center for Clinical Investigation, Yale New Haven Hospital (UL1TR001863); Georgia Clinical and Translational Science Alliance, Emory University (UL1TR002378); Weill Cornell Medicine Clinical and Translational Science Center, Weill Medical College of Cornell University (UL1TR002384); Institute for Clinical and Translational Research at Einstein and Montefiore, Montefiore Medical Center (UL1TR002556); Clinical and Translational Science Institute of Southeast Wisconsin, Medical College of Wisconsin (UL1TR001436); University of New Mexico Clinical and Translational Science Center, University of New Mexico Health Sciences Center (UL1TR001449); Clinical and Translational Science Institute at Children’s National, George Washington University (UL1TR001876); Spectrum: The Stanford Center for Clinical and Translational Research and Education, Stanford University (UL1TR003142); Indiana Clinical and Translational Science Institute, Regenstrief Institute (UL1TR002529); Center for Clinical and Translational Science and Training, Cincinnati Children’s Hospital Medical Center (UL1TR001425); Boston University Clinical and Translational Science Institute, Boston University Medical Campus (UL1TR001430); Clinical and Translational Science Institute, The State University of New York at Buffalo (UL1TR001412); Wisconsin Network For Health Research, Aurora Health Care (UL1TR002373); Advance Clinical Translational Research (Advance-CTR), Brown University (U54GM115677); New Jersey Alliance for Clinical and Translational Science, Rutgers, The State University of New Jersey (UL1TR003017); The Institute for Translational Medicine, Loyola University Chicago (UL1TR002389); Langone Health’s Clinical and Translational Science Institute (UL1TR001445); Institute for Translational Medicine and Therapeutics, Children’s Hospital of Philadelphia (UL1TR001878); Frontiers: University of Kansas Clinical and Translational Science Institute, University of Kansas Medical Center (UL1TR002366); Harvard Catalyst, Massachusetts General Brigham (UL1TR002541); ConduITS Institute for Translational Sciences, Icahn School of Medicine at Mount Sinai (UL1TR001433); Louisiana Clinical and Translational Science (LA CaTS) Center, Ochsner Medical Center (U54GM104940); HonorHealth (Voluntary); The UC Irvine Institute for Clinical and Translational Science (ICTS), University of California Irvine (UL1TR001414); Altman Clinical and Translational Research Institute, University of California
San Diego (UL1TR003107); UCDavis Health Clinical and Translational Science Center, University of California Davis (UL1TR001860); UCSF Clinical and Translational Science Institute, University of California San Francisco (UL1TR001872); UCLA Clinical Translational Science Institute, University of California Los Angeles (UL1TR001881); Northern New England Clinical & Translational Research (NNE-CTR) Network, University of Vermont (U54GM115516); UAMS Translational Research Institute, Arkansas Children’s Hospital (UL1TR003107).

Acknowledgments

Authorship was determined using the International Committee of Medical Journal Editors recommendations. The authors thank the Children’s Tumor Foundation Clinical Care Advisory Board for their support and input during the conceptualization of this project; Tim Bergquist, Yao Yan, and Justin Guinney at Sage Bionetworks for their helpful input during the design of this project; and Harold Lehmann (Johns Hopkins School of Medicine) and Kenneth Wilkins (National Institute of Diabetes and Digestive and Kidney Diseases) for helpful and insightful discussions regarding the statistical analysis strategies described in this study. The analyses described in this publication were conducted using the data or tools accessed through the National Center for Advancing Translational Sciences (NCATS) National COVID Cohort Collaborative (N3C) Data Enclave (https://covid.cd2h.org) and the N3C Attribution and Publication Policy v 1.2-2020-08-25b supported by NCATS U24 TR002306. This research was possible because of the patients whose information is included within the data and the organizations (https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories) and scientists who have contributed to the ongoing development of this community resource (https://doi.org/10.1093/jama/ocaa196). The N3C data transfer to NCATS is performed under a Johns Hopkins University Reliance Protocol # IRB00249128 or individual site agreements with National Institutes of Health (NIH). The N3C Data Enclave is managed under the authority of the NCATS; information can be found at https://ncats.nih.gov/n3c/resources. The study was supported by funding from Neurofibromatosis Therapeutic Acceleration Program (NTAP).

Author Information

Conceptualization: J.B., J.M.F., L.J.K., K.H.Y., J.T.J., S.R.P., R.J.A., J.O.B.; Data Curation: J.B.; Formal Analysis: J.B.; Funding Acquisition: J.B., R.J.A., J.O.B.; Investigation: J.B., R.J.A.; Methodology: J.B., J.M.F., R.J.A.; Project Administration: J.B., R.J.A., J.O.B.; Resources: J.B., R.J.A.; Software: J.B.; Supervision: R.J.A., J.O.B.; Validation: J.B.; Visualization: J.B.; Writing-original draft: J.B., J.M.F., R.J.A., J.O.B.; Writing-review and editing: J.B., J.M.F., L.J.K., K.H.Y., J.T.J., S.R.P., R.J.A., J.O.B.

Ethics Declaration

Data access and analysis in this study were compliant with the research protocol (Sage Bionetworks #2021101002) approved by the Western Institutional Review Board (IRB) - Copernicus Group IRB (WCG IRB) and granted IRB-exempt status. Subsequently, a request (#RP-DD0EDC, “Investigating COVID-19 burden in neurofibromatosis type 1 patients”) for access to the de-identified N3C data set (phenotypic acquisition v3.3) was submitted and approved by the N3C Enclave Data Access Committee. Clinical institutions with formally executed data transfer agreements with the National Center for Advancing Translational Sciences (https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories) contributed de-identified and retrospective electronic health record data associated with a limited set of patients relevant to COVID-19. As the steward of the data, National Center for Advancing Translational Sciences (NCATS) oversees the use of the data, and the Federal Risk and Authorization Management Program (FedRAMP) certified N3C Data Enclave through user registration, federated login, data use agreements with institutions, and data use requests with users. Thus, the data used in the study did not require informed consent from individual patients. This study does not include any individual patient information, and all analyses used aggregated and de-identified information.

Conflict of Interest

J.T.J. is a recipient of royalties from Elsevier and consulting fees from Navio Theragnostics, Inc; CereXis; and Recur- sion. S.P. is a cofounder of NFlection Therapeutics and NF2 Therapeutics, Inc; consults for Akous; and serves on the Scientific Advisory Board of and holds stock in SonALA- sense. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.jgim.2022.10.007) contains supplementary material, which is available to authorized users.

References

1. Evans DG, Howard E, Giblin C, et al. Birth incidence and prevalence of tumor-prone syndromes: estimates from a UK family genetic register service. Am J Med Genet A. 2010;152A(2):327-332. http://doi.org/10.1002/ajmg.a.33139
2. Evans DG, Bowers NL, Tobi S, et al. Schwannomatosis: a genetic and epidemiological study. *J Neurol Neurosurg Psychiatry*. 2018;89(11):1215-1219. http://doi.org/10.1136/jnnp-2018-318538

3. Uusitalo E, Rantanen M, Kallionpää RA, et al. Distinctive cancer associations in patients with neurofibromatosis type 1. *J Clin Oncol*. 2016;34(17):1978-1986. http://doi.org/10.1200/JCO.2015.65.3576

4. Rasmussen SA, Yang Q, Friedman JM. Mortality in neurofibromatosis 1: an analysis using U.S. death certificates. *Am J Hum Genet*. 2001;68(5):1110-1118. http://doi.org/10.1086/350211

5. Radtke HB, Klein-Tasman BP, Merker VL, et al. The impact of the COVID-19 pandemic on neurofibromatosis clinical care and research. *Orphanet J Rare Dis*. 2021;16(1):61. http://doi.org/10.1186/s13027-021-01911-w

6. Wolters PL, Reda S, Martin S, et al. Impact of the coronavirus pandemic on mental health and health care in adults with neurofibromatosis patient perspectives from an online survey. *Am J Med Genet A*. 2022;188(1):71-82. http://doi.org/10.1002/ajmg.a.62490

7. Haendel MA, Chute CG, Bennett TD, et al. The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. *J Am Med Inform Assoc*. 2021;28(3):427-443. http://doi.org/10.1093/jamia/ocoa196

8. Bennett TD, Moffitt RA, Hajasjo JG, et al. Clinical characterization and prediction of clinical severity of SARS-CoV-2 infection among US adults using data from the US National COVID Cohort Collaborative. *JAMA Netw Open*. 2021;4(7):e2116901. http://doi.org/10.1001/jama-networkopen.2021.16901

9. National-COVID-Cohort-Collaborative. Phenotype_Data_Acquisition. GitHub: Accessed August 11, 2021. https://github.com/National-COVID-Cohort-Collaborative/Phenotype_Data_Acquisition

10. Tolliver S, Smith ZI, Silverberg N. The genetics and diagnosis of pediatric neurocutaneous disorders: neurofibromatosis and tuberous sclerosis complex. *Clin Dermatol*. 2022;40(4):374-382. http://doi.org/10.1016/j.cldermatol.2022.02.010

11. Stone WL, Basit H, Los E. Fragile X syndrome. In: *StatPearls [Internet]*. StatPearls Publishing: 2022.

12. Becker JC, Stang A, DeCaprio JA, et al. Merkel cell carcinoma. *Nat Rev Dis Primers*. 2017;3:17077. http://doi.org/10.1038/nrdp.2017.77

13. Pelcovits A, Niroula R. Acute myeloid leukemia: a review. *Orphanet J Rare Dis*. 2013;8:35. http://doi.org/10.1016/j.orpha.2013.07.013

14. Centers for Disease Control and Prevention. United States Cancer Statistics. Incidence and death rates. Centers for Disease Control and Prevention. Updated June 6, 2022. Accessed June 15, 2021. https://seer.cancer.gov/seerstat/tutorials/aarates/dec2020

15. Anderson RN, Rosenberg HM. Age standardization of death rates: 19. Marcadis AR, Davies L, Marti JL, Morris LGT. Racial disparities in cancer presentation and outcomes: the contribution of overdiagnosis. *J Natl Cancer Inst*. 2020;102(23):20173-20179. http://doi.org/10.1093/jnci/djaa262

16. Kristal E, Pode-Shakked B, Hazan G, et al. The effects of the COVID-19 pandemic on patients with lysosomal storage disorders in Israel. *Orphanet J Rare Dis*. 2021;16(1):379. http://doi.org/10.1186/s13023-021-0079-7

17. Coombs BS, Biernacka JM. Application of the parametric bootstrap for gene-set analysis of gene-environment interactions. *Eur J Hum Genet*. 2018;26(11):1679-1686. http://doi.org/10.1038/s41431-018-0236-x

18. Považan M, Mikkelsen M, Berrington A, et al. Comparison of multi-vendor single-voxel MR spectroscopy data acquired in healthy brain at 26 sites. *Radiology*. 2020;295(1):171-180. http://doi.org/10.1148/radiol.202020191037

19. Halekoh U, Højsgaard S. A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest. *J Stat Softw*. 2014;59:1-21. http://doi.org/10.18637/jss.v059.i09

20. Pirriacchio R, Resche-Rigon M, Chevret S, Journios D. Do simple screening statistical tools help to detect reporting bias? *Ann Intensive Care*. 2013;3(1):29. http://doi.org/10.1186/2110-5820-3-29

21. Pasman E, Vidaud M, Vidaud D, Wolkenstein P. Neurofibromatosis type 1: from genotype to phenotype. *J Med Genet*. 2012;49(8):483-489. http://doi.org/10.1136/jmedgenet-2012-100978

22. Poyhonen M, Kytölä S, Leisti J. Epidemiology of neurofibromatosis type 1 (NF1) in northern Finland. *J Med Genet*. 2000;37(8):632-636. http://doi.org/10.1136/jmg.37.8.632

23. Kallionpää RA, Uusitalo E, Leppävirta J, Pöyhönen M, Peltonen S, Peltonen J. Prevalence of neurofibromatosis type 1 in the Finnish population. *Genet Med*. 2018;20(9):1082-1086. http://doi.org/10.1038/s41436-018-0086-5

24. WHO Working Group on the Clinical Characterisation and Management of COVID-19 infection. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis*. 2020;20(8):e192-e197. Published correction appears in *Lancet Infect Dis*. 2020;20(10):e250. https://doi.org/10.1016/S1473-3099(20)30483-7

25. Gregory JM, Slaughter JG, Dufuus SH, et al. COVID-19 severity is tripled in the diabetes community: a prospective analysis of the pandemic’s impact in type 1 and type 2 diabetes. *Diabetes Care*. 2021;44(2):526-532. Published correction appears in *Diabetes Care*. 2022;45(5):1299. https://doi.org/10.2337/db21-2260

26. Lim S, Bae JH, Kwon HS, Nauck MA. COVID-19 and diabetes mellitus: from pathophysiology to clinical management. *Nat Rev Endocrinol*. 2021;17(1):11-30. http://doi.org/10.1038/s41574-020-00435-4

27. Brandt JS, Hill J, Reddy A, et al. Epidemiology of coronavirus disease 2019 in pregnancy: risk factors and associations with maternal and neonatal outcomes. *Am J Obstet Gynecol*. 2021;224(4):389.e1-389.e9. http://doi.org/10.1016/j.ajog.2020.09.043

28. Richardsen S, Hirsch JS, Narasingha R. COVID-19 severity is tripled in the diabetes community: a prospective analysis of the pandemic’s impact in type 1 and type 2 diabetes. *Diabetes Care*. 2021;44(2):526-532. Published correction appears in *Diabetes Care*. 2022;45(5):1299. https://doi.org/10.2337/db21-2260

29. Burron E, Bakhai C, Kar P, et al. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *Lancet Diabetes Endocrinol*. 2020;8(10):813-822. http://doi.org/10.1016/S2213-8587(20)30272-2

30. Kristal E, Pode-Shakked B, Hazan G, et al. The effects of the COVID-19 pandemic on patients with lysosomal storage disorders in Israel. *Orphanet J Rare Dis*. 2021;16(1):379. http://doi.org/10.1186/s13023-021-0079-7