Journal of Primary Care & Community Health
Volume 12: 1–6

DOI: 10.1177/21501327211010991
journals.sagepub.com/home/jpc
SAGE

A Midwest COVID-19 Cohort for the Evaluation of Multimorbidity and Adverse Outcomes from COVID-19

Sanjeev Nanda[1], Loren Toussaint[2], Ann Vincent[1], Karen M. Fischer[1], Ryan Hurt[1], Darrell R. Schroeder[1], Audry S. Chacin Suarez[1], Jose R. Medina Inojosa[1], John C. O'Horo[1], Ramona S. DeJesus[1], Haitham S. Abu Lebdeh[1], Manpreet S. Mundi[1], Salma Iftikhar[1], and Ivana T. Croghan[1]

Abstract

Objective: To describe the process and outcome of creating a patient cohort in the early stages of the COVID-19 pandemic in order to better understand the process of and predict the outcomes of COVID-19. **Patients and Methods:** A total of 1169 adults aged 18 years of age or older who tested positive in Mayo Clinic Rochester or the Mayo Clinic Midwest Health System between January 1 and May 23 of 2020. **Results:** Patients were on average 43.9 years of age and 50.7% were female. Most patients were white (69.0%), and Blacks (23.4%) and Asians (5.8%) were also represented in larger numbers. Hispanics represented 16.3% of the sample. Just under half of patients were married (48.4%). Common comorbid conditions included: cardiovascular diseases (25.1%), dyslipidemia (16.0%), diabetes mellitus (11.2%), chronic obstructive pulmonary disease (6.6%), asthma (7.5%), and cancer (5.1%). All other comorbid conditions were less the 5% in prevalence. Data on 3 comorbidity indices are also available including the: DHHS multi-morbidity score, Charlson Comorbidity Index, and Mayo Clinic COVID-19 Risk Factor Score. **Conclusion:** In addition to managing the ever raging pandemic and growing death rates, it is equally important that we develop adequate resources for the investigation and understanding of COVID-19-related predictors and outcomes.

Keywords

pandemic, covid-19, registry, comorbidity, patient cohort

Dates received 11 January 2021; revised 25 March 2021; accepted 27 March 2021.

Introduction

Patients databases are an important part of healthcare.[1-3] Professionals and researchers in healthcare rely on well-characterized patient samples for the development of new research studies and the growth of understanding of illness and disease, especially in terms of specific or rare diseases and conditions.[1-3] For instance, the National Institutes of Health (NIH) maintains a noncomprehensive list of almost 75 active patient registries that facilitate research and discovery in areas such as Alzheimer's, heart, and kidney disease, as well as, cancer, infertility, and genetic disorders, to name just a few.[1]

The COVID-19 pandemic has shined a light on just how important patient databases can be to the development and knowledge of novel diseases. With major knowledge gaps regarding COVID-19, it is paramount to medical research that adequately characterized patient cohorts are available for COVID-19 studies of predictors, outcomes, and treatments to proceed.[4] In the 9 months since the onset of COVID-19, numerous patient databases have been developed and utilized to address important research questions. For instance, some databases have focused on acquisition of patient data on cardiovascular comorbidities,[5] while others have emphasized comorbid conditions such as diabetes[6] and cancer.[7] These are important databases, but they lack the data necessary to offer medical researchers the ability to

[1]Mayo Clinic, Rochester, MN, USA
[2]Luther College, Decorah, IA, USA

Corresponding Author:
Sanjeev Nanda, Department of Medicine, Division of General Internal Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
Email: nanda.sanjeev@mayo.edu

examine multiple comorbid conditions or diseases. The present paper describes the development of a well-characterized cohort of COVID-19 patients from the Midwest who presented for care to the Mayo Clinic in the earliest months of the pandemic.

Method

The retrospective data collection design for this database was reviewed by The Mayo Clinic Institutional Review Board (IRB) and determined to be exempt under section 45 CFR 46.101, item 2. During the study, all significant changes to study design and procedures were appropriately filed, reviewed, and approved by the IRB.

We used the EPIC platform to extract data for this database. The EPIC platform provides a robust real-time platform for Research at Mayo Clinic that combines diverse data across multiple databases and presents that data to applications as if it were single-sourced. This platform enables reproducible and accessible research across the enterprise, and it is maintained by institutional resources. Any research conducted with this database must be under the oversight of our Institutional Review Board and as such can only be accessed by investigators within any site or satellite associated with our Institution. Accessibility to this specific Mayo Clinic sub-cohort will be considered by the corresponding author at a reasonable request and following institutional guidelines.

This COVID-19 cohort database includes 1169 adults aged 18 years of age or older who tested positive in Mayo Clinic Rochester or the Mayo Clinic Midwest Health System. In this paper we describe clinical characteristics of the sample of consecutive patients who tested positive between January 1 and May 23 of 2020. All patients had a COVID-19 positive result on reverse-transcriptase–polymerase-chain-reaction (RT-PCR) assays from nasopharyngeal swab specimens.

Data abstracted from the electronic medical record (EMR) for the above population included demographic data, clinical data and comorbidity data. Patient demographic data included: age, gender, ethnicity, race, and marital status; Patient clinical data included: dates of admission and discharge, ICU admission, hospital mortality, COVID-19 mediations, height, weight, and body mass index, pregnancy status, and smoking history; Three separate assessments of comorbidity were included in the patient data: (1) Charlson Index Score;[8,9] (2) The U.S Department of Health and Human Services (DHHS) multi-morbidity score[8,10]; and (3) the Mayo Clinic COVID-19 Risk Factors Score. The Charlson Comorbidity Index is a severity weighted measure of total comorbidity burden which predicts mortality risk and includes 17 comorbidities. A score of 0 on the Charlson Comorbidity Index translates to a 98% chance of 10-year survival, while a score of 7 or greater yields a 0% chance of 10-year survival. Second, the U.S Department of Health and Human Services (DHHS) multi-morbidity score is a composite score representing the cumulative increased burden of 20 conditions (see Table 1). International Classification of Diseases, 9th Revision (ICD-9) and 10th Revision (ICD-10) codes were used to define the 20 chronic conditions.[11] Data for patients in this databases were retrieved electronically from the Rochester Epidemiology Project (REP)[12] within a 5-year capture frame from COVID-19 diagnosis. The list of the 20 conditions is provided in Table 1. Third,

The COVID-19 risk score was designed to be a Mayo specific modification to an Epic generated COVID score. It was created by Infectious diseases as a clinical tool to assist clinicians with the timely triage and care of patients who tested positive for COVID-19 in the Mayo Clinic and Mayo Clinic Health System. The goal was to identity patients with a positive test who may require a higher level of care and triage them appropriately to prevent or mitigate an adverse outcome. The score took into consideration pertinent comorbidities (hypertension, heart disease, chronic lung conditions including COPD and asthma, chronic liver or kidney disease, diabetes, immune-compromising states, and nursing home residency) in addition to age, gender, and pregnancy. Score of 0 to 2 was considered low risk, score of 3 to 5 was considered moderate risk and score over 6 was considered high risk. The validation of this COVID-19 Risk Index Score is currently in submission.[13] Patient baseline clinical characteristics (including those that comprise reported risk scores), vital signs and laboratory values were retrieved electronically using the EPIC platform, within 30 days of the COVID-19 diagnosis. Clinical diagnosis and the Charlson comorbidity index were determined based on administrative billing data using the International Classification of Diseases 9th to 10th revision. In order to decrease false positive results, 2 occurrences of a code (either the same code or 2 different codes within the code set for a given disease) separated by more than 30 days and occurring within 5 years before the index date were required for diagnosis, an approach that has been extensively validated in the past.[14,15] This information was reviewed for internal validation in duplicate by 2 independent investigators (AC and JMI) who were blinded to the patients other characteristics.

Baseline patient characteristics are summarized using mean ± standard deviation, median and range for continuous variables and frequencies counts and percentages for categorical variables.

Results

COVID-19 patients included in this sample totaled 1169. Descriptive characteristics of this patient databases are

Table 1. Twenty Chronic Conditions Selected by the U.S. Department of Health and Human Services (DHHS).

1.	Hypertension	Hypertension/high blood pressure
2.	Congestive heart failure	Congestive heart failure
3.	Coronary artery disease	Coronary artery disease Coronary heart disease Ischemic heart disease
4.	Cardiac arrhythmias	Cardiac arrhythmias
5.	Hyperlipidemia	Hyperlipidemia
6.	Stroke	Stroke Cerebrovascular disease (stroke or transient ischemic attack)
7.	Arthritis	Arthritis
8.	Asthma	Asthma
9.	Autism spectrum disorder	Autism
10.	Cancer	Cancer (all except non-melanoma skin)
11.	Chronic kidney disease	Chronic kidney disease
12.	Chronic obstructive pulmonary disease	Chronic obstructive pulmonary disease
13.	Dementia	Dementia (including Alzheimer's and other senile dementias)
14.	Depression	Depression
15.	Diabetes	Diabetes (all non-gestational)
16.	Hepatitis	Hepatitis
17.	HIV	Human immunodeficiency virus (HIV)
18.	Osteoporosis	Osteoporosis
19.	Schizophrenia	Schizophrenia
20.	Substance abuse disorders	Substance use (drug and alcohol)

included in Table 2. Patients were 18 to 99 years of age (M = 43.9, SD = 17.5). The sex distribution was 50.7% female and 49.2% male. The largest racial group was Whites (69.0%) followed by Blacks (23.4%) and Asian Americans (5.8%), and Native Americans, Native Hawaiians, and non-disclosing each represented less than 1% of the patient sample. Most patients were non-Hispanic (77.0%), with the balance being Hispanic (16.3%) and non-disclosing patients (6.7%). Just under half (48.4%) were married, less than 2% were pregnant, and the majority had never smoked (72.3%). At the time of this database creation, few Americans were able to access therapeutic medications for Covid-19 and this is reflected in that with less than 1% of patients actively using Hydroxychloroquine, Remdesivir, or Tocilizumab. Among the 41.5% patients classified as obese (body mass index (BMI) equal or greater than 30 kg/m^2) 48.9% were classified as severely obese (BMI equal to or greater than 40 kg/m^2). In terms of comorbid conditions, a total of 14 conditions are present in the database. Just over a 4th of the sample (25.1%) had a cardiovascular condition, and dyslipidemia (16.0%) and diabetes mellitus (11.2%) were also common as were chronic obstructive pulmonary disease (6.6%), asthma (7.5%), and cancer (5.1%). All other comorbid conditions were less the 5% in prevalence. Comorbidity indices showed the full range of observed scores. On the DHHS Multi-Morbidity Score 19.1% of the samples had 3 or more comorbidities, on the Charlson Comorbidity Index the average was 0 (observed range = 0-13) (The age-adjusted Charlson comorbidity index observed range was 0-16), and on the Mayo Clinic COVID-19 Risk Factors Score the average was 1 (observed range = 0-9).

Discussion

The purpose of the present work was to create a COVID-19 patient cohort database that could be used to investigate questions pertaining to risk factors for adverse COVID-19 outcomes, especially as risk pertains to comorbid conditions. COVID-19 patient cohorts remain highly relevant to understanding the development, progression, and outcomes of COVID-19, and importantly, to identifying those at highest risk of adverse outcomes.[5-7]As the COVID-19 pandemic rages on and it once again becomes the leading cause of death in older Americans,[16] and even adds considerably to excess death in younger Americans[17] the need to identify those most vulnerable so we can do everything in our power to protect them from adverse outcomes is ever more important.

While there are a few studies reporting community COVID-19 cohorts in early stages of the COVID-19 pandemic, there are no reported cohorts in the mid to late stages of the pandemic or comparison studies of cohorts before and after the COVID-19 vaccine. In a retrospective multi center cohort of hospitalized patients with COVID-19 from March 1 to April 17, 2020, Imam et al report that advanced age and increasing number of comorbidities were independent predictors of inpatient modality for COVID-19

Table 2. Characteristics of Patients.

Demographic information and characteristics	No. (%) (N = 1169)
Age, mean (SD) [range], years	43.9 (17.6) [18.0-99.0]
Gender[a]	
Female	593 (50.7)
Male	575 (49.2)
Chose not to disclose	1 (0.1)
Race[a]	
American Indian/Alaskan Native	6 (0.5)
Asian	68 (5.8)
Black or African American	274 (23.4)
Native Hawaii/Pacific Islander	3 (0.3)
White	807 (69.0)
Choose not to disclose	11 (0.9)
Ethnicity	
Hispanic or Latino	191 (16.3)
Not Hispanic or Latino	900 (77.0)
Choose not to disclose	78 (6.7)
Married[a]	565 (48.4)
Currently pregnant	20 (1.7)
Never smoker	845 (72.3)
Active COVID-19 medication	
Hydroxychloroquine	3 (0.3)
Remdesivir	3 (0.3)
Tocilizumab	1 (0.1)
Comorbidities at baseline[b]	
Total No.	1169
Hypertension	204 (17.5)
Coronary artery disease	59 (5.0)
Congestive heart failure	30 (2.6)
Dyslipidemia	187 (16.0)
Diabetes mellitus	131 (11.2)
Cerebrovascular disease	17 (1.5)
Chronic obstructive pulmonary disease	77 (6.6)
Asthma	88 (7.5)
Chronic kidney disease	45 (3.8)
Chronic liver disease	44 (3.8)
Cancer	60 (5.1)
HIV	1 (0.1)
Dementia	35 (3.0)
Substance abuse disorder[c]	14 (1.2)
DHHS multi-morbidity score, median [range][d]	1 [0-12]
None comorbidity	566 (48.4)
≥3 comorbidities	224 (19.1)
Charlson comorbidity index score, median [range][e]	0 [0-13]
COVID-19 risk factor score, median [range][f]	1 [0-9]
BMI[g], mean (SD) [range], kg/mt^2	29.8 (7.3) [15.5-66.8]
Obesity (BMI ≥ 30)[g]	427 (41.5)
Severe obesity (BMI ≥35)[g]	209 (48.9)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); COVID-19, coronavirus disease 2019; HIV, human immunodeficiency virus; SD, standard deviation.
[a]Race, ethnicity and marital status data were collected by self-report in pre-specified fixed categories.
[b]Comorbidities listed here are defined as medical diagnoses included in medical history by ICD-9 coding. These include, but are not limited to, those presented in the table.
[c]Substance abuse disorders: drug and alcohol abuse.
[d]The DHHS multi-morbidity score reflects the cumulative increase burden of 20 conditions proposed by the U.S. Department of Health and Human Services (DHHS).
[e]Charlson Comorbidity Index predicts the 10-year mortality for a patient based on age and number of serious comorbid conditions. Scores are summed to provide a total score to predict mortality. The lowest score of 0 corresponds to a 98% estimated 10-year survival rate.
[f]COVID-19 Risk Factors Score by Mayo Clinic asses identify the number of risk factors for severe illness in COVID-19 patients. A score of 6 or more reflects high risk for severe illness.
[g]Among the 1169 cohort patients, 142 patients had missing weight and height, therefore, only 1027 were included in the BMI analysis.

patient's.[18] Of the 1305 patient's in this study, mean age was 61 years (compared to 43.9 years in our sample), 53.8% of patients were male (compare to 49.2% in our sample), 66.1% of patients were African American (compared to 23.4% in our sample). The top 3 comorbidities were hypertension (56.2% in their sample, 17.5% in our sample), followed by diabetes mellitus (30.1% in their sample compared 11.2% in our sample), chronic kidney disease (17.5% in their sample compared to 3.8% in our sample). Although some of these differences are likely attribute ability to older males and African American race, it highlights the significance of multi morbidity in COVID-19 patients.

In another study describing the 1st 100 cases of COVID-19 in a hospital from Madrid with a 2 month follow-up, Munoz et al report a high mortality (23%) at 2 month follow up mostly attributable to advanced age and presence of multi morbidity.[19] The top 3 comorbidities in their sample were hypertension, heart disease and diabetes mellitus. Their adjusted charlson comorbidity index was 2 compared to out score of 0 indicating a sicker population that likely contributed to the higher mortality.

Although no longitudinal assessments utilizing this cohort have been completed to date, we speculate that this database could have several uses in the future. For example, there is currently a deficit of information on post covid syndrome, a condition characterized by fatigue, breathlessness, arthralgias, and neuropsychological symptoms.[20] One

unique aspect of this dataset is its ability to be linked to the electronic medical record. This would allow us to measure real time clinical encounters related to post covid syndrome and long term outcomes including associated morbidity and disability longitudinally in this patient cohort. This is important information as medical institutions globally prepare to provide clinical care for this little known condition. In addition, this database which represents patients with COVID-19 earlier on in the pandemic could be expanded to include patients in the middle and latter part of the pandemic. This would allow for comparative studies that provides datas on how therapeutics and vaccines altered mortablity and impacted post covid sequelae. Additionally the exact mechanism for the association of COVID-19 with cardiovascular comorbidities such as hypertension, coronary artery disease and diabetes is currently unknown. A well characterized cohort would allow future mechanistic studies related to COVID-19.

Our cohort compilation has several limitations. First, our cohort reflects the demographic characteristics of the region in which this medical center is located; therefore our sample is predominantly Caucasian and our findings therefore have limited generalizability to cohorts that are more racially and ethnically diverse. This limitation could be easily addressed by the inclusion of COVID -19 patients from 2 other states (pending IRB approval) where this medical center has a presence (AZ and FL) where the population is more racially and ethnically diverse. Second, this this database is also limited in that it consists of positive Covid-19 cases in the first phase of the Covid-19 pandemic (March 2020-May 2020) and does not contain further follow up past May 2020. This limitation too can be addressed, since the database can be added to and long term follow up through the patient electronic medical record can be obtained. Additionally the database can easily adapted to include long term follow up and it can be compared with positive cases mid to late pandemic (May-August 2020 and September-December 2020, as well as January-June 2021) to detect and trace changes in pandemic patient profiles in order to improve our knowledge on treatment and prevention.

Conclusion

This paper describes the construction of a well-characterized patient cohort established in the earliest phases of the pandemic and developed to aid in the investigation of risk factors for adverse outcomes of COVID-19. The primary advantages of this patient cohort are its early development and considerable assessment and prevalence of comorbid conditions that will aid in investigating questions of risk and vulnerability to COVID-19-related adverse outcomes. The cohort is large, well-constructed, and offers the opportunity to examine important and urgent questions regarding outcomes of COVID-19. In the future, this cohort would potentially allow for comparisons between patients afflicted by COVID-19 in earlier versus middle and later stages of the pandemic.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: All authors declare no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; and no other relationships or activities that could appear to have influenced the submitted work. RTH reports consulting fees from Nestlé, and research funding from an IIIR grant with Zealand Pharmaceuticals.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported in part by Mayo Clinic Department of Medicine and CCaTS COVID-19 Grant Award.

Ethics and Consent to Participate

In accordance with the Declaration of Helsinki, this study was reviewed and approved (ID 20-002783) by the Mayo Clinic Institutional Review Board (IRB). Mayo Clinic IRB approved informed consent waiver.

Ethical Standards

This study was determined to be EXEMPT under 45 CFR 46.101, item 2 by the Mayo Clinic Institutional Review Board which had ethical oversight for this study. In addition, the authors assert that all procedures contributing to this work comply with the ethical standards of the Mayo Clinic Institutional Review Board guidelines on human experimentation in accordance with the Declaration of Helsinki of 1975, as revised in 2008.

ORCID iDs

Loren Toussaint https://orcid.org/0000-0001-8876-1848
Ivana T. Croghan https://orcid.org/0000-0003-3464-3525

Availability of Data and Materials

All data supporting the study findings are contained within this manuscript

References

1. National Institutes of Health. List of Registries. Updated January 6, 2021. Accessed January 6, 2021. https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries
2. Gliklich RE, Dreyer NA, Leavy MB. *Registries for Evaluating Patient Outcomes: a User's Guide*. Government Printing Office; 2014.

3. Workman TA. Engaging patients in information sharing and data collection: the role of patient-powered registries and research networks. Report no.: AHRQ 13-EHC124-EF. Rockville, MD: Agency for Healthcare Research and Quality (US); 2013.
4. Alger HM, Williams JHt, Walchok JG, Bolles M, Fonarow GC, Rutan C. Role of data registries in the time of COVID-19. *Circ Cardiovasc Qual Outcomes*. 2020;13:e006766. doi:10.1161/CIRCOUTCOMES.120.006766
5. Alger HM, Rutan C, Williams JH, et al. American Heart Association COVID-19 CVD registry powered by get with the guidelines. *Circ Cardiovasc Qual Outcomes*. 2020;13:e006967. doi:10.1161/CIRCOUTCOMES.120.006967
6. Ebekozien OA, Noor N, Gallagher MP, Alonso GT. Type 1 Diabetes and COVID-19: Preliminary Findings From a Multicenter Surveillance Study in the U.S. *Diabetes Care*. Aug 2020;43(8):e83-e85. doi:10.2337/dc20-1088
7. Caruso-Davis MK, Guillot TS, Podichetty VK, et al. Efficacy of low-level laser therapy for body contouring and spot fat reduction. *Obes Surg*. Jun 2011;21(6):722-9. doi:10.1007/s11695-010-0126-y
8. Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. *Prev Chronic Dis*. 2013;10:E66. doi:10.5888/pcd10.120239
9. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45:613-619. doi:10.1016/0895-4356(92)90133-8
10. Multiple chronic conditions—a strategic framework: optimum health and quality of life for individuals with multiple chronic conditions. Accessed January 6, 2020. https://www.hhs.gov/sites/default/files/ash/initiatives/mcc/mcc_framework.pdf. Washington, DC
11. Commission on Professional and Hospital Activities. *H-ICDA, Hospital Adaptation of ICDA*. Vol. 2, The University Of Michigan; 1973
12. St Sauver JL, Grossardt BR, Finney Rutten LJ, et al. Rochester epidemiology project data exploration portal. *Prev Chronic Dis*. 2018;15:E42. doi:10.5888/pcd15.170242
13. O'Horo J. *Comparative Analysis of Covid-19 Risk Models in a Large Health System Cohort*. Personal Communication; 2021.
14. Rocca WA, Boyd CM, Grossardt BR, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. *Mayo Clin Proc*. 2014;89:1336-1349. doi:10.1016/j.mayocp.2014.07.010
15. Chamberlain AM, St Sauver JL, Gerber Y, et al. Multimorbidity in heart failure: a community perspective. *Am J Med*. 2015;128:38-45. doi:10.1016/j.amjmed.2014.08.024
16. Woolf SH, Chapman DA, Lee JH. COVID-19 as the leading cause of death in the United States. *JAMA*. 2020;325:123-124. doi:10.1001/jama.2020.24865
17. Faust JS, Krumholz HM, Du C, et al. All-cause excess mortality and COVID-19-related mortality among US adults aged 25-44 years, march-july 2020. *JAMA*. 2021;325:785-787. doi:10.1001/jama.2020.24243
18. Imam Z, Odish F, Gill I, et al. Older age and comorbidity are independent mortality predictors in a large cohort of 1305 COVID-19 patients in Michigan, United States. *J Intern Med*. 2020;288:469-476. doi:10.1111/joim.13119
19. Munoz P, Galar A, Catalan P, et al. The first 100 cases of COVID-19 in a hospital in Madrid with a 2-month follow-up. *Rev Esp Quimioter*. 2020;33:369-378. doi:10.37201/req/072.2020
20. Halpin SJ, McIvor C, Whyatt G, et al. Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: a cross-sectional evaluation. *J Med Virol*. 2021;93:1013-1022. doi:10.1002/jmv.26368