Human Subjects Protocol

Ann Arbor VA Healthcare System IRB

COVID-19 Observational Research Collaboratory

Funding Agency: VA HSR&D

Principal Investigator: Dr. Amy S. B. Bohnert, PhD, MHS

Version 7, December 16, 2024

Abstract

Background: Early evidence suggests that the frequency and type of long-term complications stemming from SARS-CoV-2 infection may depend on the severity of initial presentation. However, additional unidentified factors may also play a role including those related to the individual, the environment and/or the specific strain of the virus. The long-term health impacts of SARS-CoV-2 will have major implications for health care utilization, care processes and quality and costs of care for Veterans, both within the VA and in the community. Over 200,000 Veterans have had COVID-19. Although the long-term consequences of this novel illness are unknown, it is clear they will often be significant. Our overarching goal in the Long-Term Outcome Study (LTO) is to advance VA HSR&D research into the consequences and care of COVID-19 infection by developing and releasing integrated multi-modal data and results on long-term outcomes of COVID for use by VA clinical, research and operations communities. We will do so via a coordinated program of data production and analysis using the VA electronic health record (EHR) and other administrative sources (e.g., Medicare), de novo longitudinal survey, and qualitative inquiry. This work will be closely coordinated with the COVID-19 Observational Research Collaboratory (CORC) Coordinating Center (CCC) and the burgeoning community of HSR&D-funded COVID researchers.

This study aims to examine the long-term outcomes in Veterans infected and uninfected with SARS-CoV-2.

Aim 1: Using VINCI-CDW data, patients will be identified who had a SARS-CoV-2 test since February 2020. We will use a combination of traditional epidemiologic studies, and machine learning algorithms to determine the long-term, multi-systemic and functional adverse outcomes and/or syndromes related to SARS-CoV-2. Members of the research team are currently developing methods to identify new onset and exacerbations in comorbidities following SARS-CoV-2 illness in Veterans.

Aim 2: Determine Veteran-Reported Outcome Effects Using Structured Survey on domains such as self-reported symptoms, disability, and financial impacts, measured repeatedly over 36 months.

List of Abbreviations

ACEI	angiotensin-converting enzyme inhibitors
ARB	angiotensin receptor blockers
ARDS	Acute Respiratory Distress Syndrome
BMI	Body mass index
BSW	Bachelor of Social Work
CCMR	Center for Clinical Management Research
CDC	Centers for Disease Control and Prevention
CDW	Corporate Data Warehouse
CMS	Centers for Medicare & Medicaid Services
Co-I	Co-Investigator
COPD	Chronic obstructive pulmonary disease
CORC	COVID-19 Observational Research Collaboratory
CORC	COVID-19 Observational Research Collaboratory
COVID-19	Coronavirus Disease 2019
COVID-19	Coronavirus Disease 2019
CRP	C-reactive protein
CT	Computed tomography
DART	Data Access Request Tracker
EHR	Electronic Health Record
EKG	Electrocardiogram
HIPAA	Health Insurance Portability and Accountability Act
HIPAA	Health insurance portability and accountability act
HSR&D	Health Services Research & Development
ICU	Intensive Care Unit
ICU	Intensive care unit
IL-6	Interleukin 6
IRB	Institutional Review Board
JAMA	Journal of the American Medical Association
LASSO	Least Absolute Shrinkage and Selection Operator
LTO	Long-term outcomes
MD	Doctor of Medicine
ML	Machine Learning
MPH	Master of Public Health
MS	Master of Science
MSW	Master of Social Work
NHLBI	National Heart, Lung, and Blood Institute
NSAIDS	Nonsteroidal anti-inflammatory drugs
PD/PI	Project Director/Principal Investigator
PhD	Doctor of Philosophy
PT/INR	Prothrombin time and international normalized ratio
ResCU	REcovey after in HoSpital Cardiac arrest: late outcomes and Utilization
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
US	United States
VA	Veterans Affairs
VAAAHS	Veterans Affairs Ann Arbor Healthcare System
VAMC	VA Medical Center
VHA	Veterans Health Administration
VINCI-CDW	VA Informatics and Computing Infrastructure-Corporate Data Warehouse
VIReC	Veterans Information Resource Center
VISIN	Veterans Integrated Service Networks
WHODAS	World Health Organization Disability Assessment Schedule

Contents

Protocol Title: COVID-19 Observational Research Collaboratory

1. Introduction

The majority of patients infected with SARS-CoV-2 develop relatively mild, self-limiting, flu-like symptoms that do not require hospitalization. However, a small proportion develop severe viral pneumonia and hypoxemia, which may progress to respiratory failure, ARDS, cytokine release syndrome, multiorgan failure and death. These patients need hospitalization, frequently require admission to the intensive care unit, often need mechanical ventilation, and have a high mortality.

The COVID-19 pandemic has evolved so quickly that it is still unclear what proportion of infected patients experience adverse outcomes such as hospitalization, ICU admission, mechanical ventilation and death. It is also unclear why a small minority develop severe, life-threatening COVID-19 while the majority have mild or even asymptomatic presentation.

Furthermore, it is becoming apparent that many patients continue to experience symptoms and adverse effects from COVID-19 infection for a long period of time after the acute symptoms subside, a condition known as long COVID. A particular interest of this project is to evaluate the long-term manifestations and outcomes of SARS-CoV-2.[1-9] The literature so far suggests the following:

- Long-term adverse outcomes may occur not only among survivors of severe acute COVID-19 illness, but among the much larger group who were asymptomatic or minimally symptomatic at presentation.
- A variety of multi-systemic long-term complications are being described related to direct tissue invasion via ACE2 receptors (e.g. lungs, heart, kidneys, CNS), immunologic activation and/or hypercoagulability.
- Early evidence suggests that the frequency and type of long-term complications may depend on the severity of initial presentation. However, additional unidentified factors may also play a role including those related to the individual, the environment and/or the specific strain of the virus.
- Some individuals appear to have persistent symptoms (e.g., fatigue, joint pain and cognitive dysfunction) after clearing the virus ("COVID long haulers").[10,11] It is not known how prevalent this potential new syndrome might be, the timeline, extent and reversibility of clinical manifestations.
- A "post-ICU" syndrome has been described among survivors of sepsis, respiratory failure, and SARS-CoV and MERS-CoV including decrements in cognitive and physical function, psychological health, quality of life and loss of independence[12,13] but has not been well characterized in SARS-CoV-2.
- The long-term health impacts of SARS-CoV-2 will have major implications for health care utilization, care processes and quality and costs of care for Veterans, both within the VA and in the community.

For Aim 1, We propose to use national VA electronic data derived from the Corporate Data Warehouse to extract the electronic health records of all VA patients who tested positive for SARS-CoV-2 as well as appropriate control VA enrollees who were in care at the same time but did not test positive for SARS-CoV-2, to determine the rate of adverse outcomes, identify independent predictors and develop innovative, state-of-the-art prediction models of adverse outcomes, evaluate the effectiveness of pharmacotherapies on adverse outcomes and evaluate the impact of SARS-CoV-2 on healthcare utilization.

Aim 2 is a prospective survey of Veterans that has been funded by national HSR&D as a part of the COVID-19 Observational Research Collaboratory (CORC). The CORC is an integrated suite of activities includes secondary analyses of data, the development of this survey (submitted for review here), and other activities not yet progressed to the level of human subjects research and/or being led at other centers. Because the needs of Veterans continue to evolve across the epidemic, all CORC research could not be pre-planned—instead, the project had to be divided

into multiple phases. The CORC team, all the way up to HSR&D Director David Atkins, recognize that this places additional burdens on review process, and are grateful for your flexibility.

The primary scientific question of Aim 2 is “what are the effects of having gotten COVID-19 among Veteran survivors?” In this specific proposal, this question will be answered by comparing a random sample of Veteran survivors of COVID-19 to carefully matched non-COVID Veterans who had the same risk of acquiring COVID, but did not. Our primary outcomes are disability, health-related quality of life, mobility, caregiving receipt, financial hardship, and non-VA healthcare use.

The human subjects research aspects of this work will be conducted exclusively at VA Ann Arbor by approved study team members.

We wish to emphasize that while the application of these techniques to the specific population of Veterans with COVID is highly novel, the techniques themselves are not. They have been developed over the last 10 years at VA Ann Arbor in the study of patients who survived in-hospital cardiac arrest, where they have been highly informative. These techniques for identifying relevant Veterans, contacting them, seeking verbal informed consent, using a brief survey instrument, and following them up over time have all been previously reviewed in detail by AAVA’s RDC and IRB as part of IIRs: 1) ResCU I [IIR 13-079] and ResCU II [IIR 17-045], and approved and re-approved as part of continuing review.

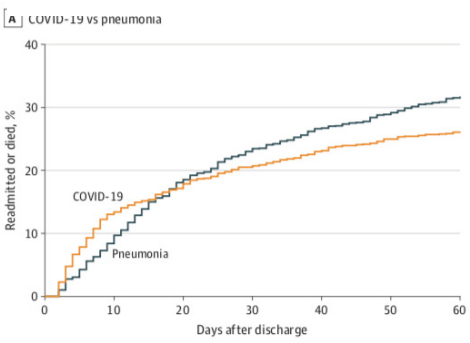
Aim 2 will utilize the VA Corporate Data Warehouse (CDW) secondary data analysis from Aim 1 to identify all Veterans to be surveyed. These data will come from the CDW and is secured under a HIPAA waiver.

Scientific Background

Scientifically, it has become increasingly clear that patients who survive the illness caused by SARS-CoV-2—that is, COVID-19—have a wide range of ongoing concerns. Whereas once some assumed the COVID-19 would be rapid but self-limited illness (ending in either mortality or recovery), it is instead clear that something is still going on with these patients. Lay press reports have termed this “Long COVID” or “Long-Haulers”, and have framed it as consistent with other post-illness syndromes or as something entirely new and different.[6-8] A robust online community has evolved (e.g. https:// www . survivorcorps . com) and the CDC has begun doing education about its existence (https:// www .cdc.gov /coronavirus/2019-ncov/long-term-effects.html). The NIH has announced a major call for research into Post Acute Sequelae of COVID.[9] By October 2020, del Rio, Collins, and Malani reviewed potential long-term consequences in JAMA and noted potential sequelae in nearly every organ.[10] The United Kingdom’s National Institute for Health and Care Excellence, the Scottish Intercollegiate Guidelines Network, and the Royal College of General Practitioners have already endeavored to release guidelines for Long-COVID, and these have been critiqued[11] as being insufficiently immunologically specific, even as that critique seems to omit any awareness of non-immunologically mediated consequences of illness.[12] No generally accepted case definition or consensus theory of “long-term outcomes” exists yet.

This may be, in part, because population-based surveillance has been woefully rare to this point, and what is out there has recreated nearly every form of bias in measurement described in any epidemiologic text—particularly highly selected samples with substantial non-random missingness in outcome ascertainment. Within these selected samples, there is often a very substantial focus on specific symptoms. 6-month follow-up from Spring 2020 survivors in Wuhan emphasized fatigue, weakness, and depression, albeit in follow-up limited to among those without readmission to the hospital or any mobility impairment.[13] A January 2021 report from Northern Italy noted frequent substantial impairments in respiratory function and mental health

Figure 1. Readmission or death after COVID-19 in VA, Spring 2020



among survivors of COVID, yet also a 65% refusal-to-participate rate.[14] Frequent worsening of mental health and limited access to care were noted among respondents to an online convenience sample of Canadians.[15]

What U.S.-based population-based data are available raise troubling signals. A multi-state follow-up of patients with a positive COVID test from spring 2020 found that a third had not returned to work 2-3 weeks later.[16] Similar rates are being seen in a new national cohort study funded by NHLBI. (unpublished data) While a report from a single home health company in New York City suggested dramatic early improvements[17], patients hospitalized at 38 hospitals in Michigan from the spring 2020 wave were followed for 60-days and found 38% were still unable to return to their normal activity, 49% reported mental health effects of their illness, 23% had lost a job because of their illness and its sequelae, and 37% reported negative financial impacts of their illness; yet this study was only able to follow-up 42% of those alive and whom it attempted to contact.[18] We recently have shown in a JAMA article[19] that Veterans with COVID-19 (n=2,179) have an increased risk of readmission or death in the 10-days but not 60-days after discharge compared to matched patients with pneumonia or heart failure. (Figure 1) Of those readmitted, the most common readmission diagnoses were COVID-19 (30.2%), sepsis (8.5%), pneumonia (3.1%), and heart failure (3.1%). During readmission, 22.6% were treated in ICU, 7.1% were mechanically ventilated, and 7.9% received vasopressors. Readmission problems after COVID were of the same order of magnitude as readmission problems after pneumonia or heart failure—well-established problems of impact.

The exponential growth of the COVID-19 literature makes a more comprehensive overview both impossible and outdated by the time it is finished. Interpreting the above key findings in the COVID-19 literature in light of existing literature on long-term outcomes of other acute infectious diseases[1,2,12,20] identifies gaps in the literature not likely to be filled by other groups, emphasizing VA HSR&D's comparative advantages.

First of all, it makes it clear that measuring the long-term outcomes of COVID-19 among Veterans is not a trivial task. Depending on the purpose, that question must be operationalized in different ways.

- For a clinician running a post-COVID clinic, the central issue is: what are the range and expected frequencies of health problems that a Veteran coming to the clinic may have?
- For those seeking to understand the natural history of COVID-19, the central issue is: what would the patient's life have been like if they had not been infected with SARS-CoV-2?
- For those studying recovery from acute illness, the central issue is: what is distinctive about COVID-19 recovery relative to other recoveries from equally severe acute illnesses?
- For those charged with planning for health systems, the central issues may be: what can be done to prevent future harm, improve the recovery of those who have already recovered from COVID-19, and plan for the ongoing healthcare needs of those still in recovery?
- For Veterans recovering from COVID-19, their families and those who care for them: what are they to expect in the future and what is the experience of having long-term complications of COVID-19 like?

The problems that patients with COVID-19 face represent the culmination of multiple different factors and dynamic processes:

- SARS-CoV-2-specific damage and recovery
- Treatments for COVID-19 and associated complications
- 2020/2021 surge-related changes in access to and quality of care
- 2020/2021 economic and social disruptions
- Their pre-existing health, access to care, and social supports and environment

Because COVID-19's distribution is non-random among Veterans—and the fundamental implications of this observation[21-23] is easy to overlook—identifying the "outcomes" of COVID-19 is not a trivial task. We propose approaches to disentangle these factors in ways relevant to the key users above. We also propose that the past literature has largely missed a focus on the distinctive issues of health services research. Past work has focused on specific symptoms and organ dysfunction, rather than taking a patient-centered approach that puts overall health needs and ability to participate in work and life at the center of the exploration. Such organ-specific science is incredibly important, but also risks missing the issues that are driving central decrements of quality of life; moreover, other groups outside of HSR&D are extensively investing in it. In many cases, not enough time has elapsed yet to understand the outcomes of COVID-19. Furthermore, past work on "big picture" recovery has overwhelmingly not had an explicit comparison group, or a comparison group matched in a relatively non-specific way. Given the non-randomness of COVID infection[21], including its fundamental racial, ethnic[24], geographic, and temporal variation[25], this means that often past work is unable to disentangle what is caused by COVID and what is caused by the other co-incident factors outlined above. We propose to correct some of these deficits.

2. Objectives

Aim 1

1. Determine the long-term multi-systemic and functional adverse outcomes of SARS-CoV-2 infection in the Veteran population
2. Determine risk factors and develop prediction models for the development of long-term adverse outcomes of SARS-CoV-2 to support risk stratification, mitigation, targeted outreach, preventive measures and treatment.
3. Evaluate the effectiveness of pharmacotherapies on the development of short- and long-term adverse outcomes of SARS-CoV-2 infection.
4. Evaluate the impact of SARS-CoV-2 on healthcare utilization, care processes and costs to estimate current and future health system needs related to COVID-19.

Aim 2

The challenge of COVID-19 is unequaled in magnitude; it is not, however, without precedent. VA clinicians have saved—and continue to save—thousands of lives by grounding their care in decades of research on viral pneumonia, sepsis, acute respiratory distress syndrome (ARDS), and infection control. While much is unknown, this proposal is informed by decades of rigorous research on the long-term consequences of acute infection,[1-5] while identifying distinctive features of COVID-19. Around Thanksgiving 2020, when we began preparing this proposal, the VA estimated approximately 100,000 Veterans had been stricken by COVID-19; that estimate as of President's Day 2021 is now over 220,000. These Veterans deserve research that is rigorous, reproducible, and relevant to inform efforts to enhance their long-term recovery. Lessons learned in the VA will benefit the wider nation. We have structured our proposal to be able to serve the Veterans already struck with COVID-19 while able to adapt to changing scientific priorities as the pandemic continues.

Our overarching goal in the Long-Term Outcome Study (LTO) is to advance VA HSR&D research into the consequences and care of COVID-19 infection by developing and releasing integrated multi-modal data and results on long-term outcomes of COVID for use by VA clinical, research and operations communities. We will do so via a coordinated program of data production and analysis using the VA electronic health record (EHR) and other administrative sources (e.g., Medicare), de novo longitudinal survey (this study), and qualitative inquiry. This work will be closely coordinated with the COVID-19 Observational Research Collaboratory (CORC) Coordinating Center (CCC) and the burgeoning community of HSR&D-funded COVID researchers. Philosophically, we believe in four co-equal metrics of success: (1) novel discoveries published in the scientific literature; (2) usefulness of our work to VA operational partners; (3) use of our work as foundation and support for other VA investigators in their own work; and (4) giving voice to Veterans experiences and challenges in recovery from COVID-19.

In this proposal, our driving scientific question throughout is: “What is the effect on Veterans of a COVID-19 infection compared to not having gotten COVID-19 at that time?” We focus on the first 36 months after COVID-19.

Aim 2, *Determine Veteran-Reported Outcome Effects Using Structured Survey*. Examine and compare a population of Veterans recovering from COVID illness with a population with no positive COVID test or treatment in the VA regarding self-reported symptoms, disability, informal caregiving receipt, and financial impacts, using a structured survey approach and linked with routinely collected administrative and clinical data. We will describe effects, moderators, and causal effects via repeated surveys.

5. Study Procedures

5.1 Study Design

Aim 1

The study design is a retrospective case control study of electronic health record data from the VINCI-CDW. The study will obtain data on all patients at VA who were tested for COVID-19 since February 2020 (from the VA COVID-19 Shared Data Resource in CDW). We will derive appropriate controls for our different analyses using data for all VA enrollees who were enrolled in VA care since February 2020 from CDW. There is no direct contact with patients.

Aim 2

We will measure clinical outcomes and sources of their variation using existing secondary data from Aim 1, and through a prospective structured survey (Aim 2). For Aim 1 and 2, we will initially conduct descriptive analyses of outcomes within Veterans with COVID-19, then subsequently define causal effects and causal moderation for subsets of pre-specified outcomes.

We have used this approach—use the VA secondary data to identify precisely eligible patients (so no screening questionnaires are asked of patients who are ultimately ineligible), followed by mail and telephone information and

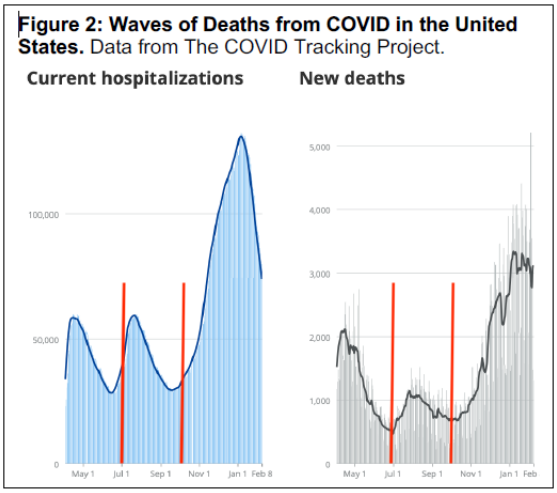
recruitment after hospital discharge—successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR's 13-079 and 17-045, approved Protocols 1597244 and 1597246).

A driving finding in VA research for the last decade has been that there is unequal and incomplete access to the benefits of VA care. EHR data by definition are only available for Veterans who have accessed care in the VA, including VA community-based care and contracted care. Other Administrative data (e.g., claims data for Medicare, Medicaid) are not available in real-time, given our dependency on other entities (i.e., CMS and VIReC) for production and linkage. Further, many important measures are not systematically ascertained or entered into the EHR in a uniform fashion for all patients. Therefore, prospective primary data collection from Veterans is necessary to define measures that are poorly captured in the VA EHR including care received at non-VA sites, caregiving receipt, disability, and employment/role function and financial toxicity, and unmet care needs in order to ensure the relevance of the work being done here. In general, we will seek to survey Veterans at 12, 18, 24, 30, and 36 months after their initial SARS-CoV-2 infection, enrolling new patients as the epidemic continues, and following patients longitudinally. The key variables of interest from new survey data for the LTO are shown in Table 2 below.

Concept	Instrument	# of items (Approx)
Activity + Quality of Life	Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL); EQ-5D-5L	18
Participation	Life Space Mobility	9+
Employment	Health and Retirement Study's Job Status Form	1
Financial Toxicity	NHLBI COVID-19 Financial Toxicity scale	10
Mental Health	PHQ-9 Depression Symptoms scale	9
Unmet Need	PROMISE instrumental support	4
Non-VA Healthcare	VIREC Utilization Measures	2
Fatigue + Symptoms	PROMIS Fatigue Short Form 7a; open-ended questions	9

Aim 2 Study Population: The population of Veterans eligible for random sampling into the survey will be the same as those eligible for inclusion in Aim 1. Therefore, all data from Aim 1 will be linkable to survey responses via a cross walk.

Cases: Our primary case definition are Veterans infected with SARS-CoV-2, as defined by the standard locally available test at the time of their infection, and identified within the COVID-NST. We will restrict to Veterans seeking primary care in VA. Viral infection index date will be defined as the date of sample collection for each

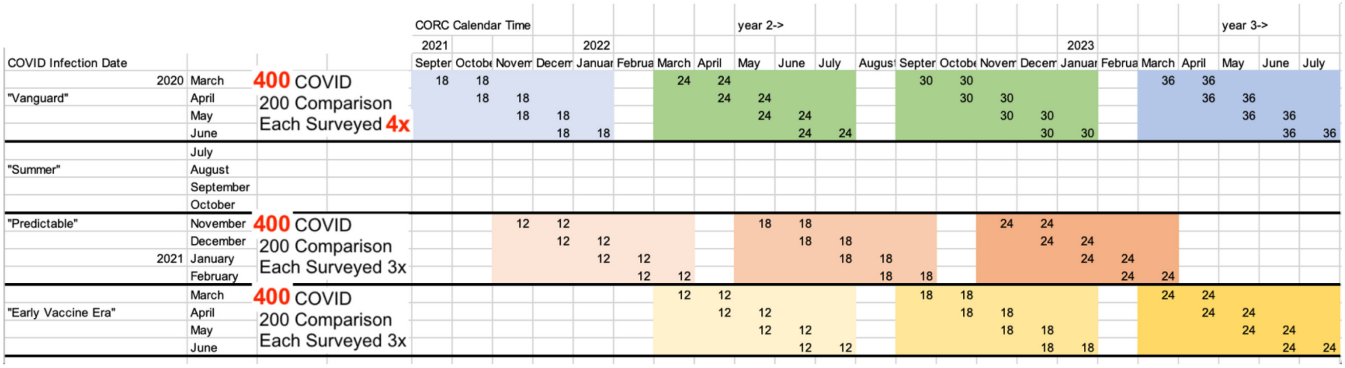


patient's first positive test. We will conduct stratified random sampling with differential survey weights so that cases are equally divided based on maximal severity of their acute COVID course in VA to include members of the following mutually exclusive groups based on patterns of health care utilization within the first 30 days after initial infection with SARS-CoV-2: not hospitalized in VA; hospitalized in VA but not mechanically ventilated; mechanically ventilated. Stratified random sampling with variable sampling weights offers two advantages: (1) sampling weights assure that the aggregated population results are still representative of the study population as a whole; while (2) ensuring adequate representation to support comparisons across these subgroups of interest. Stratum- and wave-specific sampling fractions will be chosen based on data on survival to the relevant time-point in.

In general, we will organize the sampling of cases (and later of comparators) around the waves of the epidemic, as defined by their nadirs between national death rates—each approximately 3-4 months long, as shown in Figure 2 generated based on data from the COVID tracking project. As shown in Table 3, we will initiate surveys of Veterans who have survived their initial SARS-CoV-2 infection. The viral infection "index date" will be defined as date of each patient's first positive test. This will be a longitudinal survey such that once we start surveying a Veteran, that same Veteran will be surveyed in later waves as well—we will not "resample" between, for example, the 12- and 24-month surveys.

Comparators: To support causal inference, we will also sample Veterans without COVID who are matched to participants with COVID-19. In general, cases will be compared to comparators matched on risk of developing COVID infection during a given wave, incorporating geography at the VAMC and VISN level to take into account spatiotemporal heterogeneity and small area variations. There are certain nuanced issues around precise definitions of pseudo-index dates for comparators, and particularly avoiding immortal time bias, to which we will closely attend during implementation with the input of the Methods Core. In general, we will use coarsened exact matching or propensity scores to ensure balance on key measurable covariates, both of which our group has substantial experience with.

Table 3: Example of a Timeline of Cohorts for Survey Operations. Dates on the left indicate when the COVID-19 infection occurred. Dates on the top indicate when the survey would be done. Blocks of similar surveys are shaded. The numbers 12, 18, 24, etc. represent how many months after infection each survey is, to evaluate longitudinal changes over time.



Sample Size: We will enroll up to 3,000 Veterans over the course of the survey, interviewing them each 2-4 times. This sample size is driven by power calculations needed to examine key subgroups of interest, including looking for differences across waves of COVID-19, and in racial and gender minority groups of keen interest to the VA's mission. Taking this sample size into consideration, our selection of measures emphasized continuous or multi-valued outcomes over binary outcomes to increase statistical power.

Survey operations: Our primary survey administration mode, as in past work, will be computer-assisted telephone administration by highly trained surveyors using highly validated instruments. When telephone administration is not feasible, we will offer REDCap-administered mail surveys.

We have used this approach—primarily phone-based administration of instruments by highly trained VA employees with substantial support, understanding of Veterans' issues, and of VA safety protocols—successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR's 13-079 and 17-045, Protocols 1597244 and 1597246).

We will minimize total survey error by random sampling from a group of Veterans known from the VA's CDW to ensure very high recruitment and retention rates—our group routinely exceeds recruitment rates of 80% in studies of recently critically ill patients or Veterans, with whom we have conducted over 4,000 interviews, including among individuals with substantial disability and cognitive impairment.[46] These high recruitment, retention, and survey completion rates have been achieved through careful attention to detail, stable diverse staff (some of whom are Veterans) in whom we invest in ongoing training, and with whom we engage in constant refinement of methods. These methods (including survey practices, personnel training, and customized database organization built in VA REDCap and integrated into staff calendaring and email) have been codified as Standard Operating Procedures upon which we will build.

REDCap: VA REDCap (Research Electronic Data Capture) is a web-based application hosted and maintained by the VA Informatics and Computing Infrastructure (VINCI) and VA Information Resource Center (VIReC). It is housed on a VINCI server physically located at the VA Austin Information Technology Center (AITC) in Austin, Texas and is available only on the VA network. REDCap data collection and management projects rely on a thorough study-specific data dictionary defined in an iterative self-documenting process. This includes a study-specific user rights management system administered by the Project Owner or Principal Investigator. REDCap (www.project-redcap.org) was developed specifically around HIPAA-Security guidelines. REDCap has been disseminated for use locally at other institutions and currently supports 1,700+ academic/non-profit consortium partners on six continents and over 305,000 research end-users.

Survey Measures: Our survey measure selection is driven by our conceptual agenda, ease of use, comparability with key other studies on COVID-19 and other relevant populations, and a pragmatic respect for our respondents' time and therefore the need to balance multiple competing potential measures in a sufficiently short survey. As such, we favor general purpose, brief instruments, even as we recognize the imperfections of such instruments for some specialist needs.

These survey measures are nearly all highly standard measures that have been used thousands of times in Veterans and other populations to ensure that they can be asked and answered reliably and without undue respondent burden.

Other Issues:

Performance of individual survey items will be closely monitored over time. Items will be considered for replacement if found to have significant ceiling/floor effects, disproportionate respondent burden, or diminished scientific interest as knowledge of COVID-19 evolves. Yet it is also recognized that comparability within and between cohorts is a key scientific advantage, and so core items that function well will be preserved to the extent possible.

Linkage to Aim 1 data: The scale of surveys needed to be completed, the importance of avoiding systematic non-response error, and the continuing disruption of mail service mean that written informed consent and written HIPAA authorizations are not practical. Instead, verbal informed consent and permission to link patients' survey responses with their EHR for research purposes will be obtained and recorded, allowing for full integration of survey and EHR and other sources of data. Note that this also means that we will be able to use linkages to reduce respondent burden in reporting early healthcare use.

Data Dissemination Strategy: Every six-months, beginning six months after the first survey is collected, updated releases of all structured survey items and basic summary scores will be released via VINCI for all VA investigators with appropriate permissions—but specific HIPAA-protected identifiers will not be included. The research information sheet and recruitment letter include the VA's recommended explicit language about this and the option to opt out.

Analytic Methods: Descriptive: We will summarize each of the outcomes listed in Table 2 for the patient cohort at each 6-month after the index illness. Summary statistics (mean, median, mode, range; frequency; missing data) will be reported in tabular and graphical format. Survey weights will be incorporated throughout to allow generalizability of estimates to the source Veteran population receiving VA primary care.

Analytic Methods: Variation in Survey Outcomes: A similar approach as in Aim 1 will be used for defining variation in key survey-reported outcomes, albeit with greater constraints on the number of covariates that can be included due to the more limited size of the survey sample. Multi-level models will be estimated after controlling for patient-level factors such as age, race, and summary Elixhauser comorbidity score[48], and nested within hospitals. Longitudinal growth curve models within respondents over time will be used, as we have before, to examine evolution of outcomes over time.

In general, we expect quite little missing data due to our prior experience with very high item-level completion in administering these specific instruments to Veterans after hospitalization. However, should this be a concern, we will use case-wise deletion or multiple imputation, depending on the frequency of missingness.

Analytic Methods: Causal Effects of COVID on Clinical Outcomes: Our primary causal scientific question is: “What is the effect of a COVID-19 infection compared to not having gotten a COVID-19 infection at that time on patient-reported financial and mental health functioning?”

Patients who experienced (and survived) COVID-19 will be matched to a comparator patient cohort. Causal inference is strengthened by the following strategies: (a) matching patients with and without COVID-19 on key patient characteristics, and ensuring that both groups had equal ability to be selected into the study (e.g., due to survival), (b) administering surveys on the same timeframe between COVID-19 patients and comparators, relative to a selected index date/time zero, (c) collecting sufficient data on key covariates that are unmeasured in the EHR and guided by directed acyclic graphs (DAGs) and other causal graphs generated in collaboration with the CORC’s Methods Core, (d) accounting for those covariates through inverse probability of treatment weights or similar approaches.

5.2 Recruitment Methods

Aim 1

This is a data analysis study of VINCI-CDW data. We will access the health records of all enrolled Veterans at the VA since February 2020.

Aim 2

Sample size: We anticipate 600-700 surveys per COVID-19 wave, in each of 3 waves, for a total of up to 3,000 Veterans surveyed.

Subject identification/recruitment: Our specific approach to recruitment here (after eligible Veterans are identified) is an exact replication of what we have done successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR’s 13-079 and 17-045, Protocols 1597244 and 1597246). This approach has been repeatedly reviewed and approved by VA Ann Arbor’s IRB, most recently May 2021. It is also being used in parallel studies outside VA that have been reviewed in detail by the University of Michigan and Vanderbilt University’s IRBs, acting as the central IRB for the National Institutes of Health’s COVID-19 follow-up studies.

This approach emphasizes balancing two key competing needs: (a) giving Veterans the opportunity to contribute to research and be heard; with (b) not pestering or coercing Veterans. As evidence that this approach strikes the correct balance, we note that in ResCU 1, 96% of all follow-up calls were completed. We interpret this specific data point—as well as the numerous positive comments from participants—to mean that they appreciated the opportunity to contribute. We heard over and over again that they were eager to participate in follow-up calls

because they enjoyed the first call and were glad that VA and NIH were interested in hearing their perspective. Our use of highly trained surveyors, including both Veterans and COVID-19 survivors, is crucial to providing the respectful environment that makes this possible.

Thus, we will utilize existing data from Aim 1 to identify subjects for recruitment, as outlined above. Aim 1 has been approved with a HIPAA waiver and full waiver of informed consent.

Those selected will be mailed a recruitment letter (with an option to opt out) and research information sheet describing the study, 1-2 months before the first planned survey. We will then follow-up with a phone call to discuss the study with those who do not opt out from further contact. Before each contact attempt, we will check CDW to ensure the Veteran is still alive and is not hospitalized. We will also use VHA's national electronic medical record (EMR), by way of VistaWeb and JLV.

In light of past practices and experiences, we have learned that it is important to Veterans and other patients to try a range of call times, and give them multiple opportunities to participate. In a recent study of patients recovering from respiratory failure by our group—reviewed and approved by Vanderbilt's IRB and an NIH Data Safety and Monitoring Board—it required a median and [interquartile range (IQR)] were 6 [3-13], 4 [2-10], 5 [2-11] calls to find a time when it was convenient for participants to complete the survey at 3-months, 6-months and 12-months respectively. Other studies conducted by our long-term outcomes survey team have shown similar results. In a study of patients recovering from infection, it required a median and [interquartile range (IQR)] of 2 [1-3], 3 [1-4], 2 [2-4] calls to find a time that was convenient for participants to complete the survey at 1-month, 3-months, and 6-months post-hospital discharge. In a study of patients with sepsis-induced hypotension, a median and [interquartile range (IQR)] were 4 [2-8] and 3 [2-6] calls to find a time that was convenient for participants to complete the survey at 6-months and 12-months. Again, the 80-90% repeated participation by these cohort members (as well as their comments to us) suggest this call intensity was viewed as providing flexible opportunities to participate, rather than being unwelcome. We need to be very explicit—the higher number in the interquartile range means that 1 in 4 participants who completed a survey required more than that number of calls to participate. Our analyses suggest that survivors with less education require more calls, in part because of their more variable work schedules. Setting a limit on calls too low will systematically exclude participants, particularly less educated participants, from the ability to contribute to their fellow Veterans' recovery.

Our approach to contacting Veterans who do not opt out is informed by these data, and informed by our study team members who include a COVID-19 survivor and a Veteran who has a long experience doing peer support to other Veterans. We will attempt to contact Veterans who do not opt out up to 12 times to obtain informed consent, only leaving messages 3 of those times. We will *always* cease contact at the Veterans' request, at any time; if they do not provide informed consent within the 12 contact attempts, we will also cease. If a Veteran does not respond after 7 call attempts, we will send a Consent Contact Letter to the Veteran, which will count as 1 of our 12 contact attempts. We will then make 4 more call attempts if needed. After answering all of their questions about the study, we will obtain verbal informed consent from those who agree to participate in the study and then proceed with the survey. Veterans will be reminded that they can stop the survey at any time with no adverse consequences for them. When telephone administration is not feasible, we will offer REDCap-administered web-based surveys or mail surveys.

When Veterans cannot complete the survey themselves, we will contact their proxy as indicated by the Veteran. We will follow similar recruitment procedures for the proxy (i.e. mail a recruitment letter and information sheet then follow-up via phone). Please see the Proxy Protocol for more detailed information. Any identifiable information collected for Proxy's will be stored in Access.

Participant Incentives: Our specific approach to recruitment here is an exact replication of what we have done successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR's 13-079 and 17-045, Protocols 1597244 and 1597246). Veterans will be mailed a $10 gift card for each survey to which they answer any questions at 12, 18,

24, 30, and 36 months, totaling up to $40 (Note: not all COVID-19 waves will be eligible for all surveys waves, as per Table 3 above).

5.3 Informed Consent Procedures

Aim 1

This study is a secondary analysis of data of approximately 8 million electronic VA health records. It would not be feasible to contact and subjects to obtain informed consent.

Aim 2

All local site study personnel will be trained regarding human subjects protections requirements and how to obtain and document informed consent for the aims outlined below.

- We are requesting a waiver of documentation of informed consent.
- Those selected will be mailed an information sheet describing the study, and a recruitment letter with an invitation to participate, including an opt-out option, 6 months after index illness. We will then follow-up with a phone call to discuss the study with those who do not opt out from further contact. Approved study team members will obtain verbal consent from those who agree to participate in the study and then proceed with the survey. Agreement to participant in this study will be documented in REDCap. Our specific approach to recruitment here is an exact replication of what we have done successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR's 13-079 and 17-045, Protocols 1597244 and 1597246).
- All hard copy records will be maintained in a locked cabinet in the research team's locked office. Computerized data will be stored in an electronic folder maintained on the VA Server or behind the VA Firewall, here: I:\Iwashyna CORC LTO. Only approved and trained study team members will collect and have access to study data. Only aggregate study data will be shared.

  We will ask participants if they agree to the storage, sharing, and use of their research data for future use in research studies related to COVID-19 infection and other health conditions. Researchers that access this data for future research use may be national or international and affiliated or unaffiliated with the VA, and may only use the data for research purposes. Future studies will have appropriate data use agreements in place as well as receive approval by an IRB and other oversight committees prior to starting any research. This specific approach, including the language used to inform the Veteran and of their option to opt-out, is taken directly from the VA' Cooperative Studies large "EPIC3" study of biosamples in COVID-19. We note that we will *not* have linked biosamples, yet still use their stringent approach to requesting permissions. Please see the Data Repository section on page 25 for additional details.

5.4 Inclusion/Exclusion Criteria

Inclusion:

- All VA enrollees since February 2020, which includes all patients who were tested for SARS-CoV-2 and are included in the VA COVID-19 Shared Data Resource in CDW.

Exclusion:

- VA employees

5.5 Risk/Benefit Assessment

Aim 2 Risks: This is a minimal risk study. Federal regulations at 45 CFR 46.111(a)(1) require that risks to subjects are minimized by using procedures which are consistent with sound research design. There are no study procedures, and there are no consent forms that pose a potential risk to privacy. Potential risks for this study include breach of confidentiality. All data will be collected and stored securely:

- Protected Health Information: we are requesting a HIPAA authorization waiver for the screening portion of this study. We will collect Veteran contact information to contact them for surveys and interviews, as well as social security numbers and medical record numbers. This PHI will only be kept on Microsoft Access behind the VA firewall in a protected folder, here I:\Iwashyna CORC LTO, or on VA REDCap.

- Survey Data: Survey data will not contain identifiable information. We will use REDCap to collect and temporarily store survey-based data, and other non-identifiable administrative data, such as consent date, survey date, and information related to refusal/exclusion/withdrawal. This data will be collected via the VA REDCap service by approved study team members at AAVA. We will use arbitrary study IDs to link a Veteran's identifiable information with the information we collect via REDCap. Data will be downloaded on a weekly basis to a protected folder, here I:\Iwashyna CORC LTO.

Performing this study under waiver of documentation of informed consent for Veterans will minimize risk to privacy, as there will be fewer paper records that have potential to cause a security breach.

Our efforts to protect privacy and respond to survey respondent distress are based on years of experience using specific approaches that are an exact replication of what we have done successfully in VA Ann Arbor studies including ResCU 1, ResCU 2 (IIR's 13-079 and 17-045, Protocols 1597244 and 1597246).

To minimize the risk of a breach of confidentiality, we will perform the following steps. First, as soon as the cohort is defined by the data manager, each patient in the cohort will be assigned a unique study ID. We will then create a password protected electronic tracking file that maps the study participant's identifying information to the study ID. No identifying information will be placed on data collection forms (e.g. surveys). Identifiers of potential recruits and study participants will be required to allow for follow-up contacts and to link data from different sources during the data collection phase. These will be kept in password protected files for the duration of the project and will then be destroyed according to VHA Records Control Schedule 10-1 (RCS 10-1) once direction for destruction of research records is published by VHA. Electronic data will be maintained behind a secure firewall maintained by the Ann Arbor coordinating center, here I:\Iwashyna CORC LTO. All identifiable data will be kept behind the firewall, with access only available to key study team members. The data manager will be responsible for creating analytic datasets for statisticians and investigators; these datasets will be de-identified per HIPAA guidelines. All resulting research data will be presented in aggregate only. Furthermore, study staff sign a pledge of confidentiality and understand that breach of confidentiality is grounds for dismissal. Study staff are required to complete annual training on privacy and HIPAA, as well as triennial training on human subjects protection. Access to participant data, both paper and electronic, will be given to VA project staff only and on an "as-needed" basis. There are no treatments, procedures, or interventions as part of this study. Any complaints/concerns expressed to the study staff by participants, providers, or anyone else affected by this study will be immediately reported to the PI and IRB, as will any unexpected events.

If a participant becomes distressed, the research staff will intervene as appropriate (i.e., comfort the participant and/or seek assistance from the PIs). All contacts and interviews will be conducted in a sensitive manner that protects the dignity of respondents. Research staff working with participants will be trained for situations in which a respondent expresses the need for help with a personal or emotional problem. If a respondent verbalizes this, the research team member will ask if he or she would like the contact information for a Veteran's help line. If the respondent wants such information, the staff member will provide it. In the event that a participant expresses

suicidal or homicidal ideation, or other serious psychiatric or medical symptoms, the PI will be notified immediately and advise the research staff member as to the most appropriate course of action (i.e. contacting appropriate authorities such as a mental health professional or police). Staff will follow the protocol for PHQ-9 suicidality assessment should a patient indicate suicidal thoughts while completing the PHQ-9 (see Protocol for PHQ-9 for more information). All project staff will have been trained on VA's suicide risk assessment and response guidelines, including the capability of directly connecting suicidal participants to the 24-hour VA suicide hotline.

We wish to draw attention to this approach to participant distress or suicidality because we believe it represents a particular strength of our group, both our personnel and standard operating procedures. Our team of interviewers are long-term career professionals with substantial training and experience; they include Veterans and COVID-19 survivors to insure those perspectives are baked into everything our team does. We have published our approach to safely handling disclosure of suicidal ideation in the peer reviewed literature[47], and pride ourselves on being able to ensure Veteran safety is always the highest priority.

Aim 2 Benefits: there are no direct benefits from participating in this study. Contributing to this science may build a better understanding of the long-term impacts of COVID-19 on Veterans.

5.6 Withdrawal of Subjects

For Aim 2, there are no consequences for withdrawing from this study. Subjects who wish to withdraw may do so at any time. We will keep any data collected up to the time of withdrawal. Withdrawal will be documented in REDCap.

5.7 Study Evaluations

For Aim 1, using national VA electronic health record data from VINCI-CDW, we will obtain the results of all nasopharyngeal or oropharyngeal swabs performed for SARS-CoV-2 PCR at the time of data abstraction and identify positive (detected) versus negative results.

Potential predictor variables will be extracted as of the time of testing for SARS-CoV-2 (for risk factors/prediction of hospitalization) and also as of the time of hospitalization (for risk factors/prediction of ICU admission, mechanical ventilation or death).

Using VINCI-CDW data, we will identify the following potential predictors/risk factors which are suspected to be associated with adverse outcomes in COVID-19:

Demographic factors: age, race, ethnicity, sex, socioeconomic status (including data from USVETS),

VA consolidated community care datasets (that track use of non-VA care paid for by the VA through the FEE basis file and the Patient Integrity Tool - PIT)

Case management/care coordination (VA Consult Toolbox)

Vital signs: temperature, blood pressure, heart rate and oxygen saturation, CAN score

Geographic factors: state, county, rural/urban, geocoded home address and associated ambient air pollution, distance to VA primary care and hospital

Comorbidities: diabetes, hypertension, COPD, asthma, history of solid organ transplantation, chronic kidney disease, cirrhosis, obesity (BMI), congestive heart failure, stroke, cardiovascular disease, HIV.

Severity of illness: Measured using variables used to calculate the VA case severity index.[14]

Co-infection with other viral or bacterial pathogens: influenza A or B, respiratory syncytial virus, bacterial co-infection.

Laboratory tests: CBC with differential, low lymphocyte count and high neutrophil to lymphocyte ratio (implicated in early studies as predictor), comprehensive metabolic panel [electrolytes and liver function tests], albumin (low albumin implicated as adverse predictor), brain natriuretic peptide,

PT/INR, CRP, pro-calcitonin, magnesium, troponin and IL-6 (if available, implicated as marker of cytokine storm).

Medications that may worsen outcomes: immunosuppressives, immunomodulators, corticosteroids, angiotensin converting enzyme inhibitors (ACE-I), angiotensin receptor blockers (ARB), non-steroidal anti-inflammatory drugs (NSAIDS) and thiazolidinediones which may increase the ability of the virus to enter the cells by increasing the expression of ACE2 receptors.

Supportive medications: vasopressors, inotropes, antibiotics.

Medications potentially used to treat COVID-19: remdesivir (antiviral), tocilizumab (anti-IL-6), lopinavir/ritonavir, chloroquine/hydroxychloroquine, azithromycin, ribavirin, interferon-alpha.

We will also identify receipt of imaging studies (e.g., chest X-rays and CT and MRI scans), echocardiograms, EKGs and pulmonary function tests and their finding can be extracted from the associated text integration utility (TIUs) notes in CDW

Potential long-term health outcomes we propose to extract from CDW:

Pulmonary: home O2 prescription, abnormalities in pulmonary function tests, abnormal radiographic findings on chest imaging, persistent symptoms (dyspnea, cough), development of fibrotic lung disease, bronchiectasis, and pulmonary vascular disease.

Renal: hospitalized acute kidney injury, development/worsening of chronic kidney disease, initiation of acute or long-term dialysis.

Cardiovascular: new onset or worsening hypertension, heart failure, cardiac arrest, sudden cardiac death, myocardial infarction, acute coronary syndrome, abnormal EKG or echocardiography findings, atrial fibrillation, myocarditis, hypercoagulability, DVT, pulmonary embolism and requirement for cardiac interventions (e.g., CPR, LVAD, cardioversion, coronary revascularization).

Neurological: hemorrhagic or ischemic stroke, acute neuropathy or encephalopathy, new onset seizure, rapidly progressive muscle weakness, meningoencephalitis, and acute myelitis.

Mental Health: newly diagnosed or worsening depressive disorders, depression screening, PTSD and other anxiety and mood disorders, serious mental illness (schizophrenia, bipolar, and/or psychosis), suicidal ideation, suicidal attempts, deaths due to suicide, and homelessness.

Substance Use: Presence of or treatments for unhealthy alcohol use, alcohol use disorder, tobacco use, stimulant use disorder (amphetamine and/or cocaine), opioid use disorders, and other drug use disorders (cannabis, hallucinogens and/or sedatives).

Functional Status and Pain: development of frailty or functional impairment (e.g., injurious falls, wheelchair use), cognitive impairment (dementia, delirium, memory clinic visits, prescriptions for anti-cholinergic agents) and reduced independence (e.g., inpatient rehabilitation, skilled nursing facility care, skilled home care and home-based primary care referrals) or new diagnoses of headache, joint pain, fibromyalgia, and chronic fatigue, as well as increase in number of comorbidities.

Text Integration utilities (TIU) notes: to support more details qualitative and quantitative analysis of the electronic health record, we will search text and develop natural language processing (NLP) algorithms using the VISA search tool within VINCI to study the content of documentation in the VA-wide electronic health records of cohort members. This will allow us to provide a rich description of the content of notes and create variables not available in structured data sources including information documentation related to care processes and interactions with the health care system. In order to understand the clinical context of the content of TIU notes, we will need to have access to the specialty/discipline and clinical role of the authors of clinical notes. While we will present extracts from the electronic health record in presentations and manuscripts, the names (e.g., patients, family members, clinicians, medical centers) and dates that appear in TIU notes will be removed from any materials presented or published outside the study team so that individual quotations cannot be traced to patients, providers or other individuals. In some instances, we will indicate the season (spring, summer, winter and fall) of the year that particular quotations appeared in patients' electronic health records so that quotations can be situated in the context of the evolving pandemic but will not record actual dates.

Medicare, Medicaid and USRDS linked Medicare data

We will obtain Medicare, Medicaid and USRDS linked Medicare enrollment, claims and survey data for 2018 through the most current year available through VIReC for all Veterans included in our SARS-CoV-2 treatment and comparison groups.

5.8 Data Analysis

For Aim 1, the data analysis team will consist of Dr. Berry, Dr. Green, Dr. Ioannou, Dr. Boyko and Dr. O'Hare. Dr. Green will abstract the analytic variable from VINCI-CDW. Dr. Berry, the biostatistician will work closely with Dr. Ioannou and Dr. Boyko to create the analytic models

Approach: We will use a combination of traditional epidemiologic studies and machine learning algorithms to determine the long-term, multi-systemic and functional adverse outcomes and/or syndromes related to SARS-CoV-2.

Comparison Groups: We will compare the incidence of long-term outcomes in infected and uninfected Veterans, before vs. after the pandemic and in other potentially informative control groups (e.g. untested Veterans, flu-infected patients.[15])

Unsupervised Machine Learning (ML) Algorithms: We will use ML algorithms to reveal unexpected clusters or patterns of long-term outcomes, which can then be used to generated hypothesis to be tested in more focused statistical analyses. Unsupervised ML analyses will be executed at 6-monthly sequential time points because we anticipate that patterns of long-term complications may vary depending on time since infection.

Regression Analyses: We will build multivariable models (e.g., logit, generalized linear and time to event models) to evaluate associations between SARS-CoV-2 infection and the development of specific long-term outcomes that are either suspected *a priori* (see list below) or suggested by the results of unsupervised analyses. As appropriate, we will also consider approaches such as propensity score matching and instrumental variable analysis to account for measured and unmeasured confounding, respectively.

Adverse Outcomes: We will evaluate outcomes of mortality, hospitalization, readmission, healthcare utilization and costs, days in the community as well as the organ-specific outcomes listed above.

Aim 2 data analyses are described in the study methods.

6. Reporting

As Aim 1 is a secondary analysis of data from VINCI-CDW. The only risk is a potential breach of confidentiality of privacy. This is minimized by keeping the data on password protected VINCI servers at all times. In the unlikely event of a data security incident, we will follow current facility reporting requirements: report (within 1 hour) to the study PI, VA Puget Sound Information Security Officer, Privacy Officer, R&D Associate Chief of Staff, and IRB via the director. Protocol deviations or other violations and non-compliance will be reported according to current regulatory requirements: reported to the IRB within 5 business days of becoming aware of the event.

Amy Bohnert, study Principal Investigator, will report any unanticipated problems, serious adverse events, and protocol deviations to the Ann Arbor VA IRB following local and institutional policies.

7. Privacy and Confidentiality

Aim 1

PHI will be utilized in this data analysis study. This study involves the use of VINCI-CDW data only and no direct contact with participants at any time during the study. The potential risks for harm from this research include those that may result from invasion of privacy or breach of confidentiality.

1. Data will always be maintained on VINCI-CDW. Only aggregate data without PII/PHI may be transferred from VINCI. Data will not be shared to anyone outside of the VA.
2. Analytic datasets will identify patients only by a study ID number
3. Only IRB approved study members will have access to the study data.
4. All investigators with access to data will sign the VINCI confidentiality and data use agreement forms and have many years' experience working with these data
5. Participants' protected health information will not be disclosed to anyone outside of the approved research team.
6. Standard VA data security practices will be in effect at all times, and any breach of VA data security protocol will be reported following VA Puget Sound reporting requirements (see section 6.0 Reporting).

Aim 2

Protected Health Information: we are requesting a HIPAA authorization waiver for the screening portion of this study. We will collect Veteran contact information to contact them for surveys and interviews, as well as social security numbers and medical record numbers. This PHI will only be kept on Microsoft Access behind the VA firewall in a protected folder, here I:\Iwashyna CORC LTO, or on VA REDCap.

Survey Data: Survey data will not contain PHI. This data will be collected via the VA REDCap service by approved study team members. Data will be downloaded on a weekly basis to a protected folder, here I:\Iwashyna CORC LTO.

Interview Data: Interview data may inadvertently contain PHI. This data will be collected by phone and saved on Microsoft Access behind the VA firewall in a protected folder, here I:\Iwashyna CORC LTO, or on VA REDCap.

Veteran PHI will be linked with survey and interview data via a cross walk that links their unique study ID.

We will not disclose PHI to anyone outside of this study, and then only access necessary PHI for the purposes of this study.

Standard VA data security practices will be in effect at all times, and any breach of VA data security protocol will be reported following Ann Arbor VA reporting requirements.

8. Communication Plan

The research team consisting of IRB approved study sites at VA Puget Sound (coordinating center), VA Palo Alto, VA Portland, VA Ann Arbor, and Durham VA will hold regularly scheduled meetings using Microsoft Teams. A study specific team will be set up on Microsoft Teams for the purposes of communication and secure document sharing.

9. Information Security and Data Storage/Movement

<u>VA Informatics and Computing Infrastructure (VINCI)</u>

The VA Informatics and Computing Infrastructure (VINCI) is a Department of Veterans Affairs (VA) Health Services Research & Development (HSR&D) resource center that provides a secure, central analytic platform for performing research and supporting clinical operations activities. It is a partnership between the VA Office of Information Technology (OI&T) and the Veterans Health Administration Office of Research and Development (VHA ORD). VINCI includes a cluster of servers for securely hosting suites of databases integrated from select national VA data

sources. VINCI servers for data, applications and virtual sessions are physically located at the VA Austin Information Technology Center (AITC), located in Austin, Texas. This secure enclave with 105 high-performance servers and 1.5 petabytes of high-speed data storage has multiple layers of security and disaster recovery to prevent data loss.

To ensure the protection of Veteran data, VINCI maintains compliance with the guidelines set forth by Veterans Health Administration (VHA) Handbook 1200.12, Use of Data and Data Repositories in VHA Research, and all other applicable VA and VHA policies and regulations. In addition, VINCI has undergone all security certification activities in support of obtaining an Authorization to Operate (ATO). Access to VINCI resources are approved in accordance with the requirements of National Data Systems (NDS), VHA Handbook 1200.12, Use of Data and Data Repositories in VHA Research, and all other applicable VA and VHA policies and regulations. All data transferred from VINCI is subject to audit for compliance.

VA-credentialed research staffs are granted access to study-specific data along with tools for analysis and reporting in the secure, virtual working environment through a certified VHA network computer within the VA. If not working within a VA or VHA hosted office environment containing VA network access, researchers may apply for and then access VINCI through an approved Virtual Private Network (VPN) and Remote Desktop application. The remote computing environment enables data analysis to be performed directly on VINCI servers, offering a number of advantages: uniform security standards for access; a common point of entry for all investigators who use the data; tools for analysis and reporting; tighter and more consistent control of data quality; and the ability to standardize and update terminology and format as technology and methodology improve.

Data collection

VA provides care to Veterans at over 1,400 points of care. At the core of virtually all care processes is a broadly scoped and extensively used electronic health record system known as the Veterans Information System Technology Architecture (VistA). VistA provides a longitudinal view for patients receiving care nationwide including diagnoses, procedures, medications, labs, physiologic measurements, and text notes and reports. VA uses 130 VistA implementations to provide electronic health record services nationwide for just over 20 million Veterans historically. The aggregate content of these 130 VistA systems includes 2.8 billion documents (e.g., Progress Notes, Discharge Summaries, Reports) accumulating at a rate of 865,000 each day; 7.2 billion lab values (+1.6 million each day), 4.3 billion orders (+949,000 each day), and 4.3 billion medication administrations and prescription fills (+809,000 each day).

Data are aggregated from individual VistA systems to the VA Corporate Data Warehouse where it is modeled and prepared for use. Data published by the VHA Decision Support System (DSS), Inpatient and Outpatient Medical SAS (MedSAS), VA Health Economics Resource Center (HERC) cost data, Vital Status and VA-CMS linked data files maintained by VA Information Resource Center (VIReC), CDC National Death Index VA-linked data, and several other specialty data sets can be requested through VINCI. VA National Data Services (NDS) and other data stewards regulate the right to use the data, but VINCI facilitates the process. When study requests are approved, project-specific data are extracted from source databases and placed in SQL tables accessible only to the research team and VINCI data managers.

Data storage

The study team will keep all sensitive patient data on VINCI project servers maintained by VINCI OI&T personnel and only summarized data without PHI will be downloaded from VINCI to local storage media without data steward permission. Research staff will use an audited VINCI download utility to move summarized data for reports, presentations and publications from VINCI servers to local storage media. The VINCI download utility provides an audit path including a copy of the downloaded material. All study team personnel with access to sensitive patient data will stay current on their VA approved information security training and VA approved privacy policy training. No sensitive patient data will be shared with anyone who does not have a VA appointment.

Data Access

We will use VHA's national electronic medical record (EMR), by way of VistaWeb and JLV.

The only persons having access to the data will be the PIs, the study coordinator and other persons as authorized by the PI. The PI acknowledges that he/she has the ultimate responsibility for security of research study data. OI&T personnel will extract and move study data from VA Corporate Data Warehouse (CDW) source files to the secure VINCI project space allocated to this study by VINCI. OI&T and VINCI personnel not under the purview of the PI will control the servers, network, processors, firewall and software behind the VINCI firewall, including access rights granted to study personnel.

Study data will be kept in accordance with the Department of Veterans Affairs record control schedule 10-1 (RCS 10-1). Upon completion of the research project, the PI in conjunction with the VA Information Security Officer (ISO), and in accordance with VA policy, will ensure that, study data containing sensitive, confidential information will be returned to the VA, sanitized and removed from all servers, desktops, removable storage devices, etc. When any study personnel are no longer a part of the research team, the PI will remove that person's access to all study data and notify the VA Information Security Officer of such action. VINCI personnel will be responsible for maintaining VINCI servers where study data will be kept. It will be VINCI personnel who will move, backup and remove study data from VINCI servers and who will control access to data stored on VINCI servers. The PI will request termination of data access rights for study personnel who are no longer part of the study team.

All study team personnel with access to sensitive patient data will stay current on their VA approved information security training and VA approved privacy policy training. Sensitive patient data will not be shared with anyone who does not have a VA appointment.

Data Storage Location

Study data containing sensitive patient information will be stored on VINCI servers at the Austin Information Technology Center, 1615 Woodward St., Austin, TX 78772-0001. The specific server where the data will be stored will be chosen by VINCI personnel. The server name and location within the Austin Information Technology Center may be changed at any time at the discretion of VINCI personnel.

Specialized Software

All software used to access sensitive patient data, whether provided by VINCI, or developed by the study team, will run in virtual desktop sessions on VINCI servers within the Austin Information Technology Center and behind the VINCI firewall.

Sensitive Patient Data on VINCI Servers

Sensitive patient data will not leave the VINCI environment except by explicit permission of the data steward. The data, applications and virtual computing sessions used by the study team reside behind the VINCI firewall in the Austin Information Technology Center. Researchers will use Microsoft Remote Desktop Connection software to access virtual computing sessions from their local workstation. User interface screens are generated on a local workstation but all data processing is done in a virtual computing session on a remote VINCI server. At no time in the data analysis process is data downloaded to local servers or workstations. Sensitive patient data may appear on the local screen but it is not stored on the local workstation. When the user logs out of the virtual session, no trace of the sensitive data will be available at the local workstation. Microsoft Remote Desktop Connection software uses a VA-approved secure communications protocol between the local workstation and the remote VINCI server. There is no World Wide Web access from servers behind the VINCI firewall.

All study data containing PHI or other sensitive patient data in electronic format or in paper format whether stored on VA or non-VA equipment will be maintained on password protected and encrypted computers or on VA REDCap, behind locked doors, in a secured building. No sensitive patient information will be included in any

unencrypted, character-based communication between study team members or with other VA or non-VA personnel. Study team members with access to sensitive patient data on VINCI will all have the same data access rights and restrictions.

Data Repository

VA Puget Sound (coordinating center) has received approval for the creation of a data repository based on VA electronic health records and CORC telephone survey data for investigators that will enable them to examine the short-term and long-term outcomes in Veterans infected and uninfected with SARS-CoV-2 or other appropriate controls. There are no research hypotheses associated with this repository, given its purpose is establishment of a data repository. New hypothesis will be developed by investigators, who will then be able to make use of the data placed in the repository after obtaining all necessary approvals. Survey data of those who agreed to the release of their data will be included in the repository. For full details on the repository, please see the protocol entitled "COVID-19 Observational Research Collaboratory (CORC) Research Data Repository (RDR)," created and approved by VA Puget Sound.

10. References

1. Fraser E. Long term respiratory complications of covid-19. *BMJ.* 2020;370:m3001.
2. Daher A, Balfanz P, Cornelissen C, et al. Follow up of patients with severe coronavirus disease 2019 (COVID-19): Pulmonary and extrapulmonary disease sequelae. *Respir Med.* 2020;174:106197.
3. Del Rio C, Collins LF, Malani P. Long-term Health Consequences of COVID-19. *JAMA.* 2020.
4. Carfi A, Bernabei R, Landi F, Gemelli Against C-P-ACSG. Persistent Symptoms in Patients After Acute COVID-19. *JAMA.* 2020;324(6):603-605.
5. Garrigues E, Janvier P, Kherabi Y, et al. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. *J Infect.* 2020.
6. 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.* 2020.
7. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* 2020;5(11):1265-1273.
8. Drew DA, Nguyen LH, Steves CJ, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science.* 2020;368(6497):1362-1367.
9. Lambert, N. J. & Survivor Corps. COVID-19 'Long Hauler' Symptoms Survey Report.https://dig.abclocal.go.com/wls/documents/2020/072720-wls-covid-symptom-study-doc.pdf (2020).
10. Siegelman JN. Reflections of a COVID-19 Long Hauler. *JAMA.* 2020;324(20):2031-2032.
11. Dr. Fauci Warns This COVID Syndrome is Striking Even Healthy People. Available at: https://www.eatthis.com/fauci-covid-symptoms-fatigue/ Last accessed 11/30/2020.
12. Prescott HC, Girard TD. Recovery From Severe COVID-19: Leveraging the Lessons of Survival From Sepsis. *JAMA.* 2020;324(8):739-740.
13. Fan E, Dowdy DW, Colantuoni E, et al. Physical complications in acute lung injury survivors: a two-year longitudinal prospective study. *Crit Care Med.* 2014;42(4):849-859.
14. Render ML, Deddens J, Freyberg R, et al. Veterans Affairs intensive care unit risk adjustment model: validation, updating, recalibration. *Critical care medicine.* 2008;36(4):1031-1042.
15. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology.* 2010;21(3):383-388.

16.0 Prescott HC, Angus DC. Enhancing Recovery From Sepsis: A Review. *JAMA.* 2018;319(1):62-75.

17.0 Prescott HC, Iwashyna TJ, Blackwood B, et al. Understanding and Enhancing Sepsis Survivorship. Priorities for Research and Practice. *Am J Respir Crit Care Med.* 2019;200(8):972-981.

18.0 Iwashyna TJ, Mikkelsen ME. 17.12. Medical problems in ICU survivors. In: Firth J, Colon C, Cox T, eds. *Oxford Textbook of Medicine.*2020.

19.0 Mikkelsen ME, Netzer G, Iwashyna TJ. Post-intensive care syndrome (PICS). In: Finlay G, ed.

*UpToDate*2020.
20.0 Quartin AA, Schein RM, Kett DH, Peduzzi PN. Magnitude and duration of the effect of sepsis on survival. Department of Veterans Affairs Systemic Sepsis Cooperative Studies Group. *JAMA.* 1997;277(13):1058-1063.
21.0 Yong E. Long-Haulers Are Redefining COVID-19. In. *The Atlantic*2020.
22.0 News N. 'Post intensive-care syndrome': Why some COVID-19 patients may face problems even after recovery. https://www.nbcnews.com/health/health-news/post-intensive-care-syndrome-why- some-covid-19-patients-may-n1166611. Published 2020. Accessed.
23.0 Gulliver Caspersen S. When the Doctor Is a Covid 'Long Hauler'. In. *New York Times*2020.
24.0 Collins F. Trying to Make Sense of Long COVID Syndrome. https://directorsblog.nih.gov/2021/01/19/trying-to-make-sense-of-long-covid-syndrome/. Published 2021. Accessed.
25.0 Del Rio C, Collins LF, Malani P. Long-term Health Consequences of COVID-19. *JAMA.* 2020.
26.0 Gorna R, MacDermott N, Rayner C, et al. Long COVID guidelines need to reflect lived experience. *Lancet.* 2021;397(10273):455-457.
27.0 Prescott HC, Girard TD. Recovery From Severe COVID-19: Leveraging the Lessons of Survival From Sepsis. *JAMA.* 2020;324(8):739-740.
28.0 Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet.* 2021;397(10270):220-232.
29.0 Bellan M, Soddu D, Balbo PE, et al. Respiratory and Psychophysical Sequelae Among Patients With COVID-19 Four Months After Hospital Discharge. *JAMA Netw Open.* 2021;4(1):e2036142.
30.0 Robillard R, Daros AR, Phillips JL, et al. Emerging New Psychiatric Symptoms and the Worsening of Pre-existing Mental Disorders during the COVID-19 Pandemic: A Canadian Multisite Study: Nouveaux symptomes psychiatriques emergents et deterioration des troubles mentaux preexistants pendanturantla pandemie de la COVID-19: une etude canadienne multisite. *Can J Psychiatry.* 2021:706743720986786.
31.0 Tenforde MW, Kim SS, Lindsell CJ, et al. Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(30):993-998.
32.0 Bowles KH, McDonald M, Barron Y, Kennedy E, O'Connor M, Mikkelsen M. Surviving COVID-19 After Hospital Discharge: Symptom, Functional, and Adverse Outcomes of Home Health Recipients. *Ann Intern Med.* 2020.
33.0 Chopra V, Flanders SA, O'Malley M, Malani AN, Prescott HC. Sixty-Day Outcomes Among Patients Hospitalized With COVID-19. *Ann Intern Med.* 2020.
34.0 Donnelly JP, Wang XQ, Iwashyna TJ, Prescott HC. Readmission and death after hospitalization for COVID-19 in a large multi-hospital system. *JAMA.* 2020;in press.
35.0 Wang HE, Kabeto MM, Marquita G, et al. Trajectory of Cognitive Decline After Sepsis. *Crit Care Med.* 2021;Online first.
36.0 Ioannou GN, Locke E, Green P, et al. Risk Factors for Hospitalization, Mechanical Ventilation, or Death Among 10131 US Veterans With SARS-CoV-2 Infection. *JAMA Netw Open.* 2020;3(9):e2022310.
37.0 Eastment MC, Berry K, Locke E, et al. Body mass index (BMI) and outcomes of SARS-CoV-2 among US veterans. *Obesity (Silver Spring).* 2020.
38.0 Fan VS, Dominitz JA, Eastment MC, et al. Risk Factors for testing positive for SARS-CoV-2 in a national US healthcare system. *Clin Infect Dis.* 2020.
39.0 Lopez L, 3rd, Hart LH, 3rd, Katz MH. Racial and Ethnic Health Disparities Related to COVID-19. *JAMA.* 2021.
40.0 Prescott HC, Levy MM. Survival From Severe Coronavirus Disease 2019: Is It Changing? *Crit Care Med.* 2020.

41.0 Hynes D, Edwards S, Hickok A, et al. Veterans' Use of VA Primary Care in an Era of Expanding Choice. *Med Care.* 2021;in press.
42.0 King G, Nielsen R, Coberley C, Pope J, Wells A. Comparative effectiveness of matching methods for causal inference. *unpublished manuscript.* 2011.
43.0 Iacus SM, King G, Porro G. A theory of statistical inference for matching methods in causal research. *Political Analysis.* 2019;27(1):46-68.
44.0 James PB, Wardle J, Steel A, Adams J. Post-Ebola psychosocial experiences and coping mechanisms among Ebola survivors: a systematic review. *Trop Med Int Health.* 2019;24(6):671-691.
45.0 Seminog OO, Goldacre MJ. Risk of pneumonia and pneumococcal disease in people with severe

mental illness: English record linkage studies. *Thorax.* 2013;68(2):171-176.
46.0 Yao H, Chen JH, Xu YF. Patients with mental health disorders in the COVID-19 epidemic. *Lancet Psychiatry.* 2020;7(4):e21.
47.0 Phillips KA, Morrison KR, Andersen R, Aday LA. Understanding the context of healthcare utilization: assessing environmental and provider-related variables in the behavioral model of utilization. *Health Serv Res.* 1998;33(3 Pt 1):571-596.
48.0 Gelberg L, Andersen RM, Leake BD. The Behavioral Model for Vulnerable Populations: application to medical care use and outcomes for homeless people. *Health Serv Res.* 2000;34(6):1273-1302.
49.0 Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q.* 2005;83(4):691-729.
50.0 Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying Increased Risk of Readmission and In-hospital Mortality Using Hospital Administrative Data: The AHRQ Elixhauser Comorbidity Index. *Med Care.* 2017;55(7):698-705.
51.0 Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care.* 2013;51(4):368-373.
52.0 Wagner TH, Upadhyay A, Cowgill E, et al. Risk Adjustment Tools for Learning Health Systems: A Comparison of DxCG and CMS-HCC V21. *Health Serv Res.* 2016;51(5):2002-2019.
53.0 You JH, Lee SA, Chun SY, et al. Clinical Outcomes of COVID-19 Patients with Type 2 Diabetes: A Population-Based Study in Korea. *Endocrinol Metab (Seoul).* 2020;35(4):901-908.
54.0 Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol.* 2016;183(8):758-764.
55.0 Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol.* 2016;79:70-75.
56.0 Hernan MA. Does water kill? A call for less casual causal inferences. *Ann Epidemiol.* 2016;26(10):674-680.
57.0 Schwartz S, Gatto NM, Campbell UB. Causal identification: a charge of epidemiology in danger of marginalization. *Ann Epidemiol.* 2016;26(10):669-673.
58.0 Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA.* 2010;304(16):1787-1794.
59.0 Taylor SP, Chou SH, Sierra MF, et al. Association between Adherence to Recommended Care and Outcomes for Adult Survivors of Sepsis. *Ann Am Thorac Soc.* 2020;17(1):89-97.
60.0 Watkins-Taylor C. *Remaking a Life: How Women Living with HIV/AIDS Confront Inequality.* Los Angeles: U California Press; 2019.
61.0 NHLBI PETAL Clinical Trials Network. Early Neuromuscular Blockade in the Acute Respiratory Distress Syndrome. *N Engl J Med.* 2019;380(21):1997-2008.
62.0 Bucy RA, Hanisko KA, Kamphuis LA, Nallamothu BK, Iwashyna TJ, Pfeiffer PN. Suicide Risk Management Protocol in Post-Cardiac Arrest Survivors: Development, Feasibility, and Outcomes. *Ann Am Thorac Soc.* 2017;14(3):363-367.
63.0 van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care.* 2009;47(6):626-633.
64.0 Hynes DM, Stroupe KT, Fischer MJ, et al. Comparing VA and private sector healthcare costs for end-stage renal disease. *Med Care.* 2012;50(2):161-170.
65.0 Humensky J, Carretta H, de Groot K, Brown MM, Tarlov E, Hynes DM. Service utilization of veterans dually eligible for VA and Medicare fee-for-service: 1999-2004. *Medicare Medicaid Res Rev.* 2012;2(3).
66.0 Hynes DM, Maciejewski ML, Atkins D. HSR Commentary: Linking VA and Non-VA Data to Address Important US Veteran Health Services Research Issues. *Health Serv Res.* 2018;53 Suppl 3:5133-5139.
67.0 Hynes DM, Koelling K, Stroupe K, et al. Veterans' access to and use of Medicare and Veterans Affairshealth care. *Med Care.* 2007;45(3):214-223.
68.0 Kaul B, Hynes DM, Hickok A, et al. Does Community Outsourcing Improve Timeliness of Care forVeterans With Obstructive Sleep Apnea? *Med Care.* 2021;59(2):111-117.
69.0 Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ.* 2001;20(4):461-494.
70.0 O'Neill S, Kreif N, Grieve R, Sutton M, Sekhon JS. Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Serv Outcomes Res Methodol.* 2016;16