What Drives Long-COVID?

Understanding complex interactions with real world data.

ARCADIA®

As many as 1 in 10 people may have symptoms of COVID-19 for months after their initial infection, driving a staggering total number of long-COVID cases that continues to increase. As part of our work with the COVID Patient Recovery Alliance, Arcadia analyzed a massive real world data set to identify potential drivers of long-COVID.

Our key finding is that vaccination against COVID-19 reduces the likelihood that patients will experience long-COVID symptoms in the event that they are infected. Our data also suggest that a single dose of a COVID-19 vaccine, even if received after the patient is diagnosed with COVID, still correlates with a reduced likelihood of presenting long-COVID symptoms. We've intended this summary of our analysis to be useful for a broad healthcare audience; for more detailed clinical or technical information about our COVID-19 work and research data asset, please visit **arcadia.io/c19**

Our study "Reduced Incidence of Long-COVID Symptoms Related to Administration of COVID-19 Vaccines Both Before COVID-19 Diagnosis and Up to 12 Weeks After" (Michael A. Simon, Ryan D. Luginbuhl, Richard Parker) is currently available on **MedRxiv**. →

"I bring to your attention the fact that
a number of individuals who virologically
have recovered from [their COVID-19]
infection, in fact have persistence
measured in weeks to months of
symptomatology that does not appear
to be due to persistence of the virus.
They're referred to as long haulers.
They have fatigue, myalgia, fever, and
involvement of the neurological system,
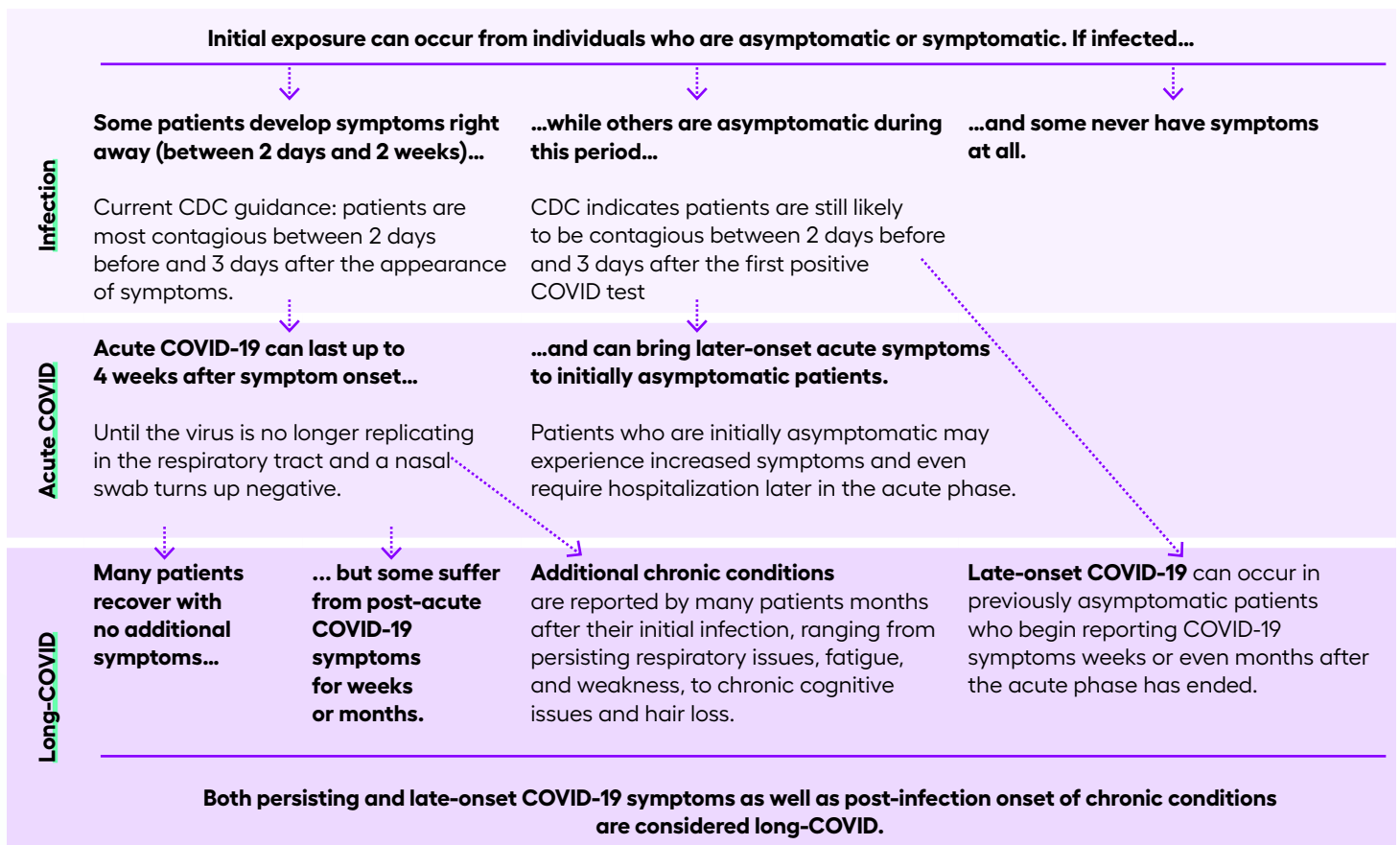as well as cognitive abnormalities,
such as the inability to concentrate."

Dr. Anthony "Tony" Fauci
Testimony to the Senate HELP Committee
September 23, 2020

Long-COVID

The COVID Patient Recovery Alliance defines long-COVID as the currently-accepted term to describe signs and symptoms that continue or develop after acute COVID-19 that are not explained by an alternative diagnosis.

It includes both ongoing symptomatic COVID-19 as well as post-COVID-19 syndrome, the onset and persistence of COVID symptoms during and after infection.

COVID symptoms can appear weeks after diagnosis

Initial exposure can occur from individuals who are asymptomatic or symptomatic. If infected...

Infection

Some patients develop symptoms right away (between 2 days and 2 weeks)...

Current CDC guidance: patients are most contagious between 2 days before and 3 days after the appearance of symptoms.

...while others are asymptomatic during this period...

CDC indicates patients are still likely to be contagious between 2 days before and 3 days after the first positive COVID test

...and some never have symptoms at all.

Acute COVID

Acute COVID-19 can last up to 4 weeks after symptom onset...

Until the virus is no longer replicating in the respiratory tract and a nasal swab turns up negative.

...and can bring later-onset acute symptoms to initially asymptomatic patients.

Patients who are initially asymptomatic may experience increased symptoms and even require hospitalization later in the acute phase.

Long-COVID

Many patients recover with no additional symptoms...

... but some suffer from post-acute COVID-19 symptoms for weeks or months.

Additional chronic conditions are reported by many patients months after their initial infection, ranging from persisting respiratory issues, fatigue, and weakness, to chronic cognitive issues and hair loss.

Late-onset COVID-19 can occur in previously asymptomatic patients who begin reporting COVID-19 symptoms weeks or even months after the acute phase has ended.

Both persisting and late-onset COVID-19 symptoms as well as post-infection onset of chronic conditions are considered long-COVID.
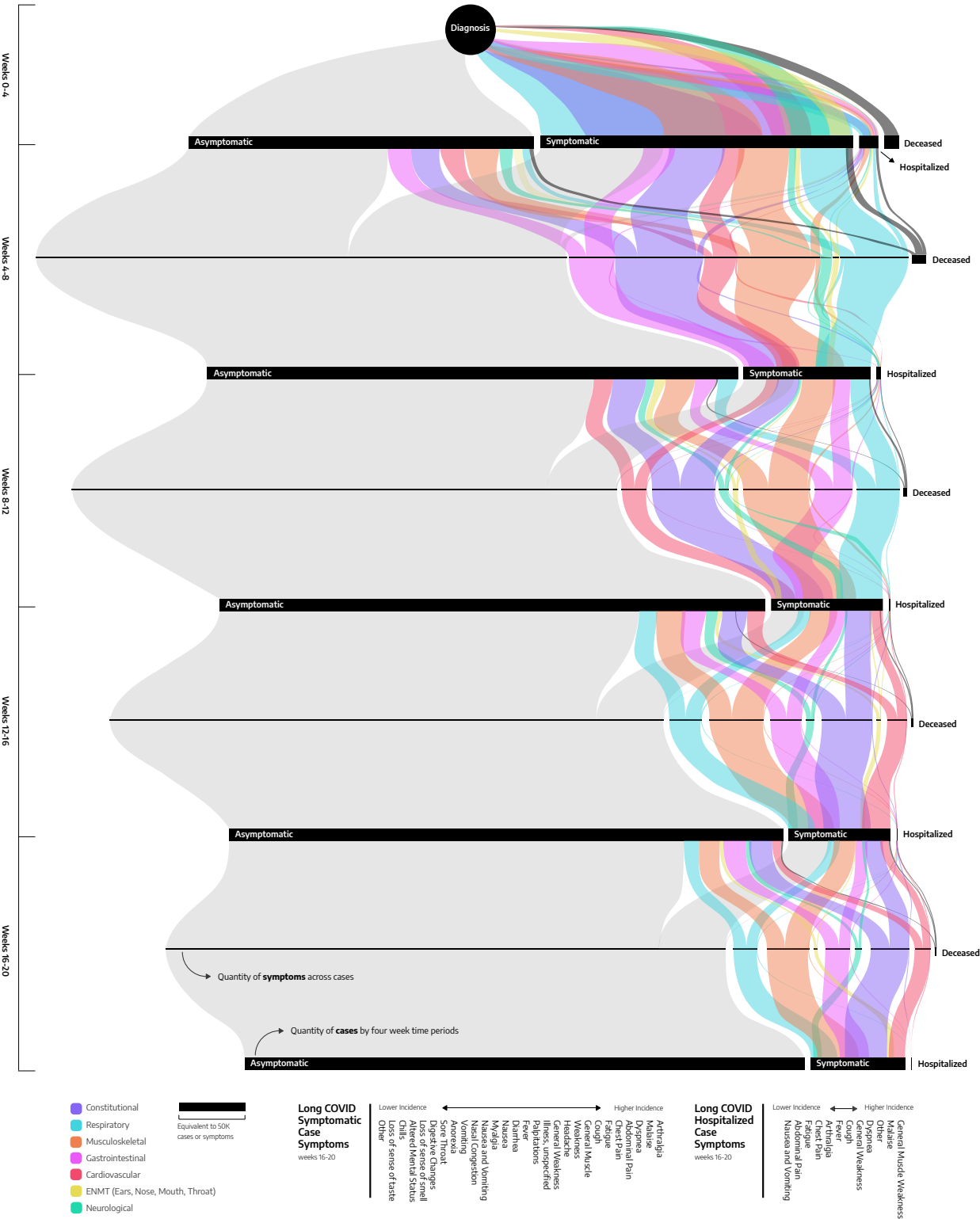

Long-COVID is a critical research priority for the COVID Patient Recovery Alliance, which brings together key thought leaders in business, health care, research, academia, data and analytics, and patient advocacy to develop national solutions that coordinate diverse data sources, inform the development of models of care, and ensure adequate payment for long-COVID patients who served their communities and nation during the pandemic; whose COVID-19- related costs are extraordinary and burdensome; or who are underserved by existing programs.

Arcadia is a contributor to the Alliance, lending our research data asset and data science expertise to this important cause.

1 Million COVID-19 cases and their symptomatic journey

What happens once a patient is diagnosed with COVID-19?
The journey varies. Some remain asymptomatic. Others develop immediate symptoms (gastrointestinal, cardiovascular, and/or others) and then quickly recover—or may be hospitalized. Some only begin to develop symptoms weeks after diagnosis. And for many, COVID symptoms persist for many weeks and may seem unending.

There is no single long-COVID experience. Patients suffer from a wide range of symptoms over highly variable timeframes. The complex nature of this condition makes it challenging to identify potential drivers—but data science can offer some insights.

View more visualizations in our online gallery.

Quantity of **symptoms** across cases

Quantity of **cases** by four week time periods

Equivalent to 50K cases or symptoms

- Constitutional
- Respiratory
- Musculoskeletal
- Gastrointestinal
- Cardiovascular
- ENMT (Ears, Nose, Mouth, Throat)
- Neurological

Long COVID Symptomatic Case Symptoms
weeks 16-20

Lower Incidence ← → Higher Incidence

Other
Loss of sense of taste
Loss of sense of smell
Digestive Changes
Altered Mental Status
Chills
Anorexia
Sore Throat
Vomiting
Nasal Congestion
Nausea and Vomiting
Nausea
Myalgia
Diarrhea
Palpitations
Fever
Illness, unspecified
General Weakness
Headache
General Muscle Weakness
Weakness
Cough
Fatigue
Chest Pain
Abdominal Pain
Dyspnea
Malaise
Arthralgia

Long COVID Hospitalized Case Symptoms
weeks 16-20

Lower Incidence ← → Higher Incidence

Other
Dyspnea
Cough
Arthralgia
Chest Pain
Fever
General Weakness
Fatigue
Abdominal Pain
Nausea and Vomiting
General Muscle Weakness
Malaise
Other

Weeks 0-4
Weeks 4-8
Weeks 8-12
Weeks 12-16
Weeks 16-20

Diagnosis

Asymptomatic
Symptomatic
Deceased
Hospitalized

4

Exploring factors driving long-COVID

Arcadia used exploratory data analysis to surface factors that might increase patient susceptibility to long-COVID symptoms. Exploratory data analysis does not assign causal relationships, but it does help understand complex interactions and drive hypothesis creation. Our findings identify strong directions for productive future research.

Our analysis involved the following practical, concrete steps, which can be applied to other problems where you need to understand potential relationships between complex interactions using a massive real world data set.

1. Analyze patient conditions
We established baselines for patients who had been diagnosed and treated for COVID-19 using a broad spectrum of chronic and acute conditions, both COVID-associated and not. Understanding the prevalence of chronic conditions in the populations we studied helped us characterize the incidence of COVID-related symptoms and conditions as patients recovered.

2. Select a data set
We needed a real-world data set that not only had a sufficient population of COVID-19 patients but also represented a more general population of active patients to support an exploratory analysis. Arcadia's research data asset includes de-identified, aggregated, HIPAA-compliant data on over 100 million patients. At the time this study was conducted, our data included 1.3 million COVID-19 positive patients of whom ~689,000 had data from an electronic medical records (EMR) system.

Arcadia's COVID-19 Population as of January 2022

2.8M	**2.4M**	**8.9M**
COVID-19 positive patients	**with data from an EMR (88%)**	**COVID-19 vaccinated patients**

■ **Patients diagnosed with COVID-19**   ■ **Patients first vaccinated against COVID-19**
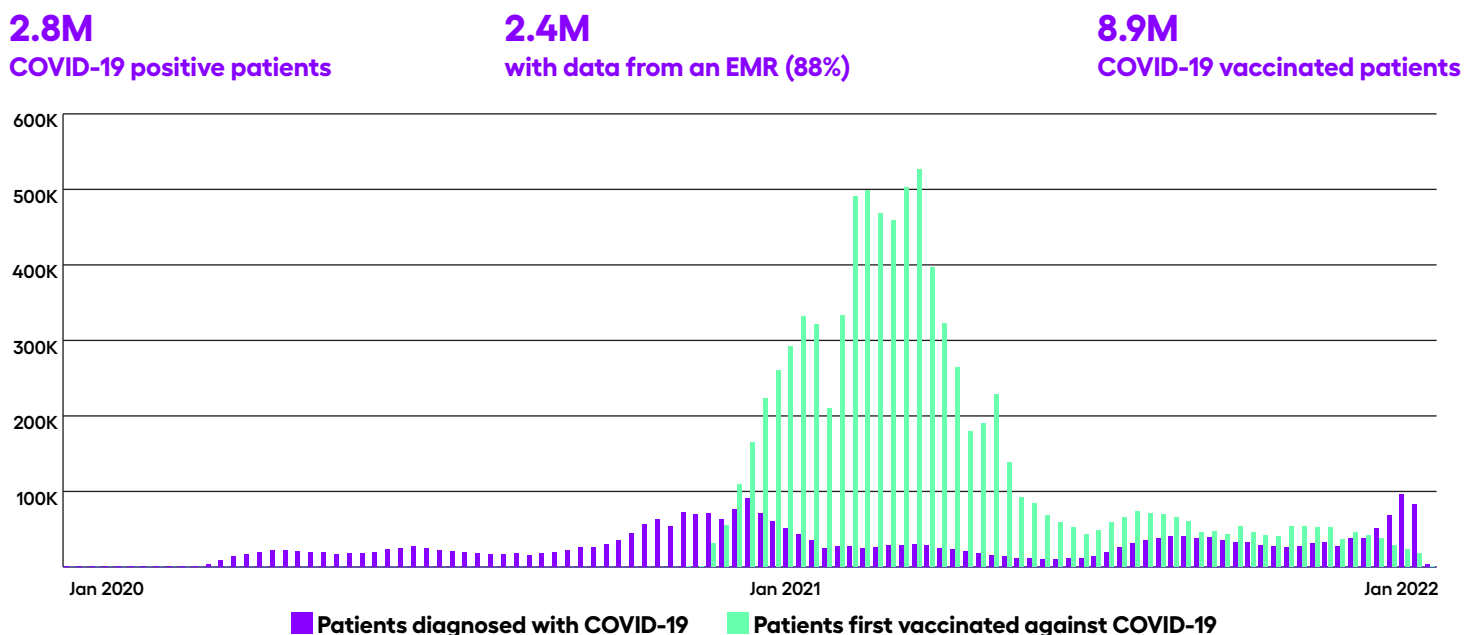
Arcadia's massive and growing real world data asset

Arcadia's research data asset was built to support the needs of healthcare organizations taking on value-based care. Insights generated from the full spectrum of healthcare data support research into both common conditions like diabetes and rare diseases like spinal muscular atrophy (SMA).

Data Asset Overview

National Coverage
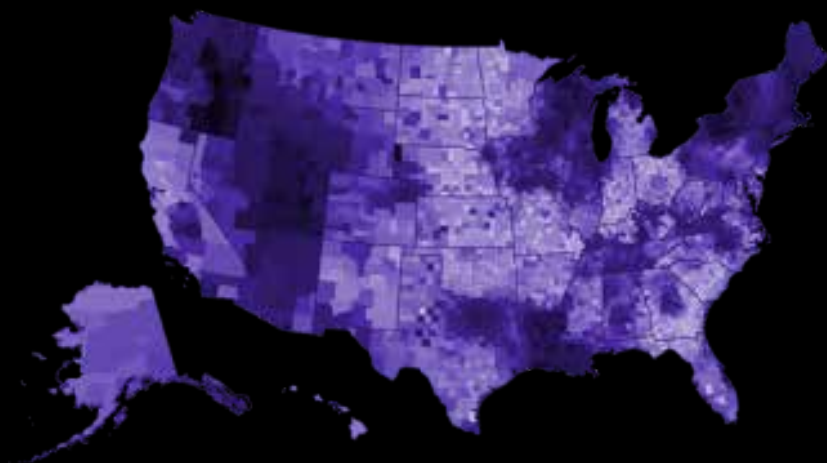
100M+
Patient lives

24%
Integrated claims records

Full Spectrum of Healthcare Data

$357B
Health costs

348M
Clinical visits

150K
Distinct EHR NPIs

2.8B
Diagnoses & problems

5.9B
Lab results

1.1B
Written prescriptions

Therapeutic Area Coverage

4.6M Diabetic Patients
Sourced from 45M active patients

3.6M
Have 2+ year history

1.9M
Have 5+ year history

1.9M
Have 3+ A1c lab results

19K
Distinct NPIs

Comorbid Condition Profile

Condition	Count
Depression	365,219
COPD	517,605
Heart	583,679
Vascular	704,444
Obesity	836,282
CHF	873,623
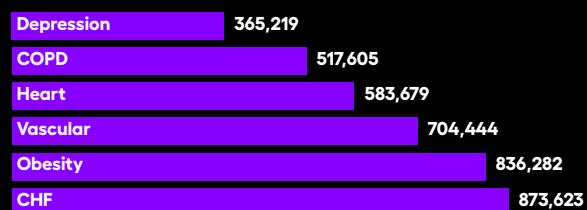
1,877 Patients with Spinal Muscular Atrophy
Sourced from 60M active patients

1,108
Diagnosed in last 2 years

24
Taking Spinraza

3. Define and refine cohorts

We included 25.8 million patients of any age who were seen clinically at least once prior to January 1, 2020 and were alive as of March 1, 2020. Of those, our COVID-positive (COV+) group included 1.06 million patients with a positive COVID-19 PCR or antigen test or a diagnosis by a provider in a clinical setting.

Both our COV+ patients and the 24.7 million patient control group (COV-) had diverse racial and ethnic representation.
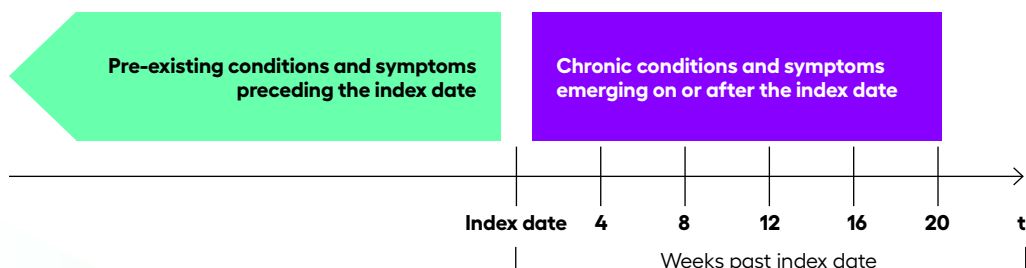
4. Classify conditions

A "condition" is a clinical diagnosis that associates a disease, sign, or symptom with a patient at a point in time. We used the AHRQ Chronic Condition Indicators to split out chronic and acute events. We also used AHRQ's clinical classification system[1] to define broad diagnostic groupings. We then identified COVID-associated conditions[2] that superseded these broad diagnostic groupings.

5. Insert events of interest

In addition to diagnostic criteria, we added other events sourced from EHR or claims-based data that we thought might be useful to our model, like hospitalizations related to a COVID diagnosis, vitals and labs, medications, and COVID-19 vaccinations.

6. Understand timing and condition onset

Looking at events the week they occurred (rather than the date) allowed us to simplify our view and compress the data without reducing their value.

---

1 More information about the AHRQ Clinical Classification Software and its classifiers is available at:
   https://www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp
2 Using value sets of COVID-associated conditions courtesy of Clinical Architecture with their permission.

7. Analyze prevalence

We looked at the prevalence of conditions present prior to our study period (and verified the data against existing literature) to create baselines for the COV+ group and the COV- control group. Here, we started to see some qualitative differences between the groups—for example, COV+ patients were 11.7% more likely to suffer from obesity.

8. Analyze incidence

As we continued to work through our analysis, we needed to figure out what separated long-COVID patients from the COV+ group and from the COV-control group.

Incidence counts indicate the onset of new or acute conditions or symptoms. By analyzing shifts in incidence counts relative to the timing of COVID diagnosis and treatment, we established a baseline. We then compared the baselines for the COV+ group against clusters that have experienced long-COVID symptoms.

Exploratory data analysis is particularly well-suited to pulling out hidden insights about complex interactions. By analyzing the symptomatic journeys of more than 1 million COVID patients, our model helped us surface factors that strongly correlate with whether a patient develops long-COVID.

Our findings

The COVID data and our initial analysis support the hypothesis that vaccination reduces the odds of experiencing long-COVID symptoms, even if the vaccination occurs after COVID-19 infection.

- Patients who are vaccinated and then infected are 4–6 times less likely to experience long-COVID symptoms and 7–10 times less likely to report more than one symptom.

- Patients who are first vaccinated after they are diagnosed with COVID-19 are still less likely to experience multiple long-COVID symptoms compared to patients who remain unvaccinated: 4-6 times less likely if vaccinated within a month after diagnosis and 3 times less likely if vaccinated within two months of diagnosis.

- The protective effect of a vaccine administered after infection declines within 2 months of infection.

Vaccines reduce the likelihood of developing long-COVID

First vaccine dose received...	Reduced likelihood of any long-COVID symptom				Reduced likelihood of more than 1 long-COVID symptom			
	Likelihood	Odds ratio	95% CI	p-value	Likelihood	Odds ratio	95% CI	p-value
Prior to diagnosis	4.5x	0.220	0.196 – 0.245	< 0.005	8.8x	0.113	0.090 – 0.143	< 0.005
0–4 weeks after diagnosis	2.6x	0.382	0.353 – 0.413		5.3x	0.189	0.163 – 0.220	
4–8 weeks after diagnosis	1.9x	0.535	0.506 – 0.567		3.2x	0.317	0.289 – 0.348	
8–12 weeks after diagnosis	1.3x	0.747	0.713 – 0.784		2.2x	0.458	0.426 – 0.493	

What are the key findings from this study?

These findings were an unexpected and exciting development from Arcadia's ongoing work with the COVID-19 Patient Recovery Alliance. The most important takeaways are:

- COVID vaccines not only reduce the incidence and terrible effects of acute COVID infection but they also reduce the likelihood and severity of long-COVID symptoms in the aftermath of that infection.

- These findings offer evidence that vaccination soon after infection is still highly protective against long-COVID, extending the window of opportunity for the unvaccinated to protect themselves against long-COVID symptoms.

- The protective effect of vaccination against long-COVID increases the personal protection offered by vaccination to the recipient, which may encourage increased prophylactic vaccination and, with continued support from evidence, therapeutic vaccination.

What else can we learn about long-COVID from a massive RWD dataset?

A massive RWD dataset like Arcadia's permits us to explore a number of additional avenues of study, including:

- The differences in long-COVID experiences based on demographic factors, including age, sex, race, ethnicity, and socioeconomic status.

- The impact of pre-existing conditions on presence and severity of long-COVID symptoms, including the onset of new chronic conditions.

- Differences in long-COVID symptoms for patients treated with different medications and therapeutics during the acute phase of their COVID infection.

- Variation in expression of long-COVID symptoms based on combinations of likely COVID variant and vaccination status.

- Impact of multiple COVID infections and vaccinations on long-COVID, especially as their effects on the immune system continue to rise in importance.

We are grateful for the opportunity to better serve our customers and contribute work to address a monumental public health crisis, and we look forward to continuing our work with the COVID-19 Patient Recovery Alliance.

For more information about our COVID-19 work,
please visit **arcadia.io/c19**

To learn more about Arcadia's RWD, please contact **hello@arcadia.io**

About Arcadia

Arcadia helps payers and providers put their data to work so they can transform healthcare. We do that through an interoperable data platform that uses advanced analytics to shape strategies, inform decisions, and facilitate actions. In turn, payers and providers can focus on what matters most — whether that's patient outcomes, operational efficiencies, or financial performance. We're trusted by the institutions driving the future of healthcare, including Southwestern Health Resources, Tandigm, Castell, Rush Health, and Beth Israel Lahey Health. To learn how Arcadia is shaping the future of healthcare with innovative solutions that deliver data-driven insights, visit **arcadia.io**.

ARCADIA®