

NEWS | 20 June 2022

How common is long COVID? Why studies give different answers

Enormous databases do not necessarily allow scientists to solve long COVID mysteries, such as how well vaccination protects against the condition.

Heidi Ledford



Health-care workers monitor a woman in a COVID-19 recovery gym in Genoa, Italy. Credit: Marco Di Lauro/Getty

Clinical epidemiologist Ziyad Al-Aly has access to a treasure trove that many researchers can only dream of: millions of sets of electronic medical records from the

US Department of Veterans Affairs (VA), which provides health care for the country's military veterans.

With this data in hand, Al-Aly, who is based at the VA St. Louis Healthcare System in Missouri, and his colleagues have delved into the long-term effects of COVID-19, from [cardiovascular illness¹](#) to [diabetes²](#). They have also undertaken the challenge of studying long COVID – a condition in which people experience symptoms months after an acute SARS-CoV-2 infection seems to have resolved – and recently published findings³ that surprised some researchers. The team found that previous vaccination only [reduces the risk of developing long COVID after infection by about 15%](#), which is [substantially less than some other estimates⁴](#), which suggested that vaccines halved the risk.

RELATED



The pandemic's true health cost: how much of our lives has COVID stolen?

It is the kind of whiplash result that people following long-COVID research have become accustomed to seeing, as data from various studies report discordant results. Differences in how the syndrome is defined, the kinds of data used to study it and how those data are analysed have left both the public and policymakers grappling with disparate answers to basic questions. How frequent is long COVID? And how does vaccination or reinfection or the latest SARS-CoV-2 variant affect the risk of developing the condition?

The answers to those questions can be used to develop COVID-19 policies, but the steady drip–drip of seesawing studies can also cause confusion. says Al-Aly. Having so much uncertainty doesn't engender a lot of trust, Al-Aly adds: "The public does not react very well to saying 'between 15% and 50%'."

Slippery definition

Part of the problem is the definition of long COVID, which has been linked to more than 200 symptoms, the severity of which can vary from inconvenient to debilitating.

The syndrome can last for months or years, and has a distressing tendency to reappear, sometimes months after an apparent recovery.

So far, there is no agreement on how to define and diagnose long COVID. The World Health Organization's attempt at a [consensus, published in 2021](#), has not proved popular with patient advocates or researchers, and studies continue to use a range of criteria to define the condition. Estimates of its prevalence can range from 5–50%.

A study of such a complex condition needs to be sufficiently large to reflect the range of symptoms and the possible impact of characteristics such as age and the severity of the acute SARS-CoV-2 infection. This is where analyses like Al-Aly's offer a host of advantages: data from large health-care networks can provide enormous sample sizes. Al-Aly's study of long COVID after a 'breakthrough' infection – one that follows vaccination – included records from more than 13 million people. Although 90% of those people were men, that still left 1.3 million women in the analysis, Al-Aly notes, more than many other studies can muster.

Big-number benefits

These large numbers, as well as the types of data available in some health records, allow researchers to perform complicated statistical analyses to carefully match the demographics of people infected with coronavirus to an uninfected control group, says Theo Vos, an epidemiologist at the Institute for Health Metrics and Evaluation at the University of Washington in Seattle, who has worked with a variety of data sources to study long COVID.

But there are also drawbacks. "People mistake the size of the study with its quality and its validity," says Walid Gellad, a physician who studies health policy at the University of Pittsburgh in Pennsylvania.

RELATED

In particular, Gellad worries that studies that rely on electronic health records will be muddied by behavioural differences. For example, compared with someone who does not seek medical care for acute



Long COVID risk falls only slightly after vaccination, huge study shows

COVID-19, someone who does might be more likely to report long-COVID symptoms, he says.

Moreover, medical records and health insurance claims might not reflect a demographically diverse population, says computational epidemiologist Maimuna Majumder at Harvard Medical School in Boston, Massachusetts. This is particularly likely in the United States, she says, where health insurance coverage varies widely. “The number of data points considered is often so large that we mistakenly assume that these data must be representative,” she says. “But this isn’t necessarily the case.”

Majumder also wonders whether studying claims data could lead researchers to undercount the number of people with long COVID, because many people might not seek medical care for their condition.

Coding lessons

Another issue is how symptoms are recorded in the claims and electronic medical records. Doctors often record codes for several symptoms and conditions, but they rarely list a code for every symptom a patient is experiencing, says Vos, and the choice of codes for a given condition might vary from one doctor to the next. This could lead to differences in whether and how long COVID is reported. “Electronic health records have useful information in them, without a doubt,” says Gellad, who says that the VA study was particularly well designed. “But for answering the question of how common something is, they may not be the best.”

Other methods also have their pitfalls. Some studies rely on self-reporting, such as the COVID Symptom Study app developed by King’s College London and the data-science company ZOE, also in London. Data from the app showed that vaccination reduced people’s risk of experiencing long COVID 28 days or more after an acute infection by about half⁴. But studies in which people voluntarily self-report their symptoms can be biased, because people who have symptoms are more likely to

participate, says Gellad. And studies that rely on smartphone apps might not fully capture data from disadvantaged communities.

RELATED



Pandemics disable people – the history lesson that policymakers ignore

One particularly useful source of data has been the UK Office for National Statistics (ONS), says Nisreen Alwan, a public-health researcher at the University of Southampton, UK. In May, the ONS reported that [the variant of SARS-CoV-2 that people are infected with can affect their risk of developing long COVID](#). Among double-vaccinated participants, those thought to have COVID-19 caused by the Omicron BA.1 variant were roughly 50% less likely to develop long COVID

symptoms four to eight weeks after infection than were participants whose infections were probably caused by the Delta variant. This finding is in line with the results of an 18 June paper⁵ based on ZOE data.

Seeking a common thread

Alwan, who has long COVID and has advocated for the collection of data on the condition, praises the ONS study design, which involved enrolling a group of people with careful attention to representing the UK population, and then following up with them to ask about their infection status and symptoms.

Other aspects of study design, such as whether a control group is used, can strongly affect results, says Alwan. But accounting for disparate methods and definitions need not stall research. “That’s not something new,” she says. “It’s something that we had before COVID, for other conditions.”

For Al-Aly, the discrepancies among study results are not surprising, nor are they damning. Epidemiologists often weave together evidence from multiple sources of data and methods of analysis, he says. Even if it is difficult to precisely quantify vaccination’s effect on long-COVID risk, for example, researchers can look for trends. “You search for the common thread,” Al-Aly says. “The common thread here is that vaccines are better than no vaccines.”

doi: <https://doi.org/10.1038/d41586-022-01702-2>

References

1. Xie, Y., Xu, E., Bowe, B. & Al-Aly, Z. *Nature Med.* **28**, 583–590 (2022).
2. Xie, Y. & Al-Aly, Z. *Lancet Diabetes Endocrinol.* [https://doi.org/10.1016/S2213-8587\(22\)00044-4](https://doi.org/10.1016/S2213-8587(22)00044-4) (2022).
3. Al-Aly, Z., Bowe, B. & Xie, Y. *Nature Med.* <https://doi.org/10.1038/s41591-022-01840-0> (2022).
4. Antonelli, M. *et al. Lancet Infect. Dis.* **22**, 43–55 (2022).
5. Antonelli, M., Pujol, J.C., Spector, T. D., Ourselin, S. & Steves, C. J. *Lancet* **399**, 2263–2264 (2022).

Latest on:

[SARS-CoV-2](#) Epidemiology Medical research

BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection

ARTICLE | 17 JUN 22



FDA authorizes COVID vaccines for the littlest kids: what the data say

NEWS | 17 JUN 22



How months-long COVID infections could seed dangerous new variants

NEWS FEATURE | 15 JUN 22

Nature (*Nature*) | ISSN 1476-4687 (online) | ISSN 0028-0836 (print)

© 2022 Springer Nature Limited