Long COVID affects twice as many Americans as official counts show, new AI study finds

Benjamin Mateus
31 May 2026

A new artificial intelligence study published last week in JAMA Network Open has found that roughly one in six Americans who contracted COVID-19 developed long COVID, more than double the rate captured by current federal surveillance. The findings, led by researchers at Mass General Brigham, lay bare a public health crisis hiding in plain sight, one systematically obscured by the very diagnostic tools that health systems and policymakers rely upon to track it.

The research team deployed a precision-phenotyping algorithm called P2RC across the electronic health records of 457,950 patients at 58 hospitals in four U.S. regions. Rather than relying on a single billing code, the standard but deeply inadequate ICD designation for post-COVID conditions, the AI read each patient's full medical history before and after infection, identifying new chronic symptoms that could not be explained by preexisting conditions. What they found was striking. Some 16.3 percent of patients met the definition for long COVID, with regional rates ranging from 13.6 percent in Western Pennsylvania to 22.7 percent in Southern California. Code-based surveillance, by comparison, captures fewer than 7 percent of cases.

"Over 10 million people with long COVID would go entirely undetected by the diagnostic code that health systems and policymakers rely on to track the disease burden," said corresponding author Hossein Estiri of Mass General Brigham. "The figures we uncovered are almost certainly an undercount." Lead author and data scientist Jiazi Tian described what that invisibility looks like from inside a clinical setting. "The cardiologist seeing new dysautonomia, the endocrinologist seeing new metabolic disease, the neurologist confronting unexplained cognitive complaints," Tian wrote, "some of these presentations are long COVID arriving without the label that would connect them to a COVID-19 infection."

Perhaps the most alarming finding is that there is no plateau in sight. Cumulative prevalence continued rising through at least mid-2024, with statistical modeling pointing to sustained growth over the coming decade if current patterns hold. Roughly 14.5 percent of infected patients developed new chronic conditions requiring sustained clinical care, among them heart disease, diabetes, cognitive impairment, fatigue syndromes, and metabolic and endocrine disorders. The researchers are explicit that long COVID should be understood as a chronic disease burden rather than a self-limited post-viral syndrome. Its systemic undercounting, they warn, impedes both epidemiologic surveillance and coordinated clinical management and may partly explain the observed post-pandemic increases in diabetes, cardiovascular disease, and fatigue syndromes.

This picture is consistent with the weight of existing science. Work by Dr. Ziyad Al-Aly and colleagues at the Department of Veterans Affairs, drawing on millions of patient records with large uninfected control groups, established clear causal links between COVID-19 and long-term organ damage, even after mild or asymptomatic infection, including elevated risks of heart disease, stroke, diabetes, kidney disease, and neurological disorders lasting years after the acute phase. The NIH RECOVER initiative and major international cohort studies converge on a long COVID prevalence of between 10 percent and 25 percent among infected adults, with a substantial subset remaining severely ill one to three years post-infection. The Brigham's 16.3 percent figure sits squarely within that consensus, and the specific organ system damage it identified, particularly profound metabolic and endocrine consequences, mirrors the chronic disease burden detailed in the Veterans Affairs research. Varying data sources and distinct methodological approaches all point in the same direction.

The burden of this disease does not fall evenly. A 2024 cohort study published in the *Annals of Internal Medicine* found that financial hardship and structural inequality significantly raise the risk of developing persistent

© World Socialist Web Site

symptoms. A 2023 U.K. study analyzing more than 200,000 working-age adults found that people in the most socioeconomically deprived areas carried substantially higher long COVID risk than those in the least deprived. The occupational picture is equally telling. Frontline healthcare workers and education workers, people who could not work from home and who were required to remain present to keep the economy functioning, faced the highest rates of the condition.

The economic consequences have compounded alongside the health toll. Harvard economist David Cutler estimated the total cost of long COVID to the U.S. economy at $3.7 trillion, a figure that rivals the aggregate economic damage of the Great Recession. Of that total, $2.2 trillion reflects lost quality of life, nearly $1 trillion comes from reduced earnings, and $528 billion represents direct medical spending. A Brookings Institution analysis found that millions of full-time-equivalent workers left the labor force entirely because of long COVID. The people hardest hit by the disease are, overwhelmingly, the same people whose labor the economy depends upon.

The invisibility of long COVID is not a measurement error. It is an institutional outcome. Flawed billing codes, fragmented care, the absence of systematic patient follow-up, and the political decision to declare the pandemic over have each played a role. The recent cancellation of NIH RECOVER program grants and the closure of the federal Office of Long COVID Research and Practice are not administrative housekeeping. They are decisions about which suffering gets counted and which gets buried. The P2RC tool is significant precisely because it demonstrates what properly structured clinical data can reveal when the will exists to look, what the Brigham's own researchers describe as a surveillance gap that public health agencies are no longer tracking.

That raises a question worth asking directly. For whom does public health exist? What P2RC demonstrates is that the answer has always been a political choice, not a technical limitation. For years, the bureaucratic underpinning of disease surveillance, a single flawed billing code, fragmented specialty care, no mechanism to follow a patient across time, functioned as a form of censorship that did not merely fail to count the sick but made them uncountable. P2RC cuts through that architecture by reading the full longitudinal record and doing what the official system was structured to prevent: connecting a COVID-19 infection to the chronic conditions that accumulate in its wake across months and years, tracing how health deteriorates progressively and persistently in ways no billing code was ever designed to capture. Although the capacity to do this has always existed, the will has not, and that gap is not a technical failing but a political one.

Essential workers, low-wage earners, people in the most economically precarious communities, those who have borne the greatest share of this disease, have the most at stake in demanding that tools of this kind are built, funded and controlled in the public interest rather than managed for administrative convenience or private profit. Scientific visibility is not the same as political action, and the distance between the two is not an oversight, but a deliberate policy.

The malign neglect applied to the initial COVID pandemic is now being applied in full force to long COVID. The systematic disappearance of this disease into flawed billing codes, its burial under the claim that the pandemic is over, the defunding of RECOVER, the closure of the federal Office of Long COVID Research and Practice, none of these are administrative failures or the result of insufficient data. They are expressions of a deliberate choice to subordinate the health of working people to the resumption of economic activity on the basis of capitalism and the protection of profit. Governments declared the pandemic over not because the science said so, but because the capitalist class required it. The Brigham study is one more demonstration of what that choice has cost, and who has been made to pay it.

The data produced by studies like this one must reach workers and their organizations. It must be used to demand safe workplaces, adequate ventilation, paid sick leave, disability recognition for long COVID, the restoration of public health infrastructure, and universal access to care provided free at the point of need. Long COVID is not a medical anomaly or a legacy problem winding its way to resolution. It is a mass chronic disease still compounding with every new wave of infection, still falling hardest on those least able to absorb it. Addressing it seriously means not only confronting who has been made to carry this burden, but why, and building the political power necessary to change it.

To contact the WSWS and the Socialist Equality Party visit:

wsws.org/contact

© World Socialist Web Site