The following methods accompany a Commonwealth Fund blog post, Eric Schneider and Arnav Shah, "Will the Pandemic Increase the Number of Americans with Preexisting Conditions?" *To the Point* (blog), Commonwealth Fund, Oct. 7, 2020.

The Commonwealth Fund
Affordable, quality health care. For everyone.

HOW WE MADE OUR PROJECTIONS

We used Centers for Disease Control and Prevention (CDC) data (see below) to calculate that 67 percent of U.S. COVID-19 cases occurred among nonelderly adults (ages 20–59) and 68 percent of the cases among nonelderly adults were not associated with a known condition. We then applied these percentages to the Johns Hopkins University database, which contains the cumulative number of COVID-19 cases (7.5 million as of October 7, 2020). We used this approach because the Johns Hopkins database lacks demographic data on cases.

The CDC's COVID-19 Case Surveillance Public Use dataset contains 4,481,062 lab-confirmed or probable COVID-19 cases (as of September 30, 2020). After removing 5,188 cases with missing age, we calculated that approximately 67 percent of reported cases involving adults ages 20 to 59. We excluded adults 60 to 64 because the CDC file reports age in decades.

The CDC Case Report form lists the underlying medical conditions and risk behaviors available for recording. After removing cases with missing data on conditions, we used the remaining data (n=1,502,309) to calculate that, among adults ages 20 to 59 with a lab-confirmed or probable COVID-19 infection, 32 percent had underlying conditions. Twenty-nine percent of cases reported no underlying medical conditions, and 39 percent of case were coded as "unknown" with regard to underlying conditions. For this analysis, we assumed that the 68 percent who reported no condition or "unknown" were otherwise healthy prior to COVID-19 infection. The CDC's COVID-19 Surveillance Public Use dataset does not include information on the individual conditions for each case.

Applying these percentages to the Johns Hopkins confirmed COVID-19 cases, we first calculated that 5 million reported cases involve nonelderly adults ages 20 to 59. Based on the proportion of the CDC surveillance cases that did not report an underlying medical condition, we estimated that just over 3.4 million nonelderly adults had COVID-19 as a new preexisting condition.

To better understand how different the population of nonelderly adults with COVID-19 might be from all nonelderly adults, we calculated within each decade of age the percentage of COVID-19 cases with at least one underlying condition. Table 1 shows the results.

TABLE 1

Among patients with lab-confirmed or probable COVID-19 case:

Age group	Share with underlying medical condition
20–29	23%
30–39	28%
40–49	36%
50–59	44%

Note: Patients with "missing" condition response, and/or "NA"/"unknown" age response, were dropped.

Data: Centers for Disease Control and Prevention, "COVID-19 Case Surveillance Public Use Data," CDC Case Surveillance Task Force, last updated Sept. 30, 2020.

We compared these estimates with a comparable calculation based on data from the Behavioral Risk Factor Surveillance System and National Health Interview Survey. See Table 2 for the results.

TABLE 2

Among all nonelderly adults:

Age group	Share with declinable condition
18–34	18%
35–44	24%
45–54	29%
55–64	44%

Data: Gary Claxton et al., *Pre-Existing Condition Prevalence for Individuals and Families* (Henry J. Kaiser Family Foundation, Oct. 2019).

While not directly comparable, the tables are consistent with one another in two ways. Both show a gradient of increasing probability of an underlying condition as age increases. They also show a similar percentage of adults having an underlying condition within the different age bands, although the COVID-19 population appears to have a higher prevalence of underlying conditions in the 20-to-39 age range compared with the national population estimates.

Our estimates have limitations. The two databases used to calculate the estimates are incomplete, even though they are the best available data on the pandemic. The Johns Hopkins database contains only confirmed cases, and it undercounts people who were infected but asymptomatic or never tested — many of whom may still go on to develop health complications.

Our estimates for the nonelderly adult population are extrapolated from a CDC data file that is incomplete with regard to the number of COVID-19 cases in the United States and is missing data on age and conditions for many cases. Because of the way CDC data are reported, we were unable to include adults ages 60 to 64 in the calculation of the proportion of nonelderly adults with preexisting conditions. It is possible we are underestimating the proportion of COVID-19-infected individuals with conditions. However, the proportions of nonelderly people with conditions based on the CDC data are higher in every age group than those calculated from the general population.