**Step 6, Estimation of Mediated Effect**

In order to calculate the mediated impact of a variable on the outcome, we calculate the total causal effect of the outcome under two networks. The first network is the best fit for the data. The second network removes the arc between the factor and the mediator. This model is often referred to as the counterfactual network as it models a world in which mediation does not occur. In our COVID example, Figure 3 indicates the best fit to the data. Note that Chills is predicted by demographic characteristics and earlier symptoms of Fever, fatigue, shivering, (please see Table 3).

**Figure 3: Observed network with a link between Fever and Chills.**



Suppose we want to understand the percent of impact of fever mediated through Chills. To create the counterfactual network, we remove the arc between fever and chills and recreate the network. This is done by regressing Chills on its predictors, but now excluding the fever variable. The original regression is in Table 3, in the column heading Chills. The modified regression without the fever variable produced the following probability of Chills:

$$prob(Chills)=\frac{e^{0.45\*Shivering+0.48\*Fatigue+0.17\*Loos of taste+0.02\*headache-0.08\*Loss of Smell-0.07\*Age-0.03\*Female+ 0.08}}{1+e^{0.45\*Shivering+0.48\*Fatigue+0.17\*Loos of taste+0.02\*headache-0.08\*Loss of Smell-0.07\*Age-0.03\*Female+ 0.08}}$$

These probabilities were made part of the parameters of the counterfactual network. These probabilities show the frequency of chills when fever is not its predictor. The modified network is presented in Figure 4 below:

**Figure 4: Counterfactual network, excluding the link between Fever and Chills**



Given these two networks, we can now calculate the unconfounded causal impact of fever on diagnosis of COVID-19. To calculate the unconfounded effect of Fever on diagnosis of COVID-19, we stratify the direct predictors of Fever. Five variables predict before fever: age, gender, loss of taste, fatigue, and shivering. These five variables create 25 = 32 combinations, shown in Table 6. For each combination, Table 6 reports how changes in fever affect probability of a positive COVID-19 test results. In Table 6, these probabilities are first calculated from Figure 3, the best fit to the data. They are shown in the portion of the Table to the right and in white. For example, in the strata (combination) where the patient is shivering, fatigued and has loss of taste, the change in fever results in .7% increase in probability of a positive COVID-19 test result. The impact of change in fever is reported for all strata. The probability of observing the strata is calculated as the product of probability of observing each of the three predictors of fever. This is shown in Table 6 in the right most column. The expected change in the probability of the outcome, shown as $∆p\left(Strata\right), $is calculated across all strata as:

$$∆p\left(∆ in X\right)=\sum\_{Strata}^{}p\left(Y\right|∆ in X, Stratum)p(Stratum)$$

This formula is known as adjustment formula as it shows how the various variables that are part of the strata adjust the value of impact of change in the X exposure on the outcome Y. In the case of fever, this equation is written as:

$$∆p\left(∆ in fever\right)=\sum\_{}^{}p\left(COVID\right|∆ in Age, Gender, fever, Shivering, Fatigue, Loss of taste)\*p(Age, Gender, Shivering, Fatigue, Loss of taste)$$

The presence of fever was calculated to increase the probability of COVID-19 by 0.19 points. Some of this impact comes from the fact that fevers can lead to chills and thus the impact of fever is mediated by chills. To understand how much of the impact of fever is mediated by chills, we recalculate the formula but now use the counterfactual network. Table 6 shows the results in the section that is colored in gray. In the counterfactual network chill is no longer mediating the effect of fever. In this network, across the strata, the change in fever increased the probability of COVID-19 by 0.14 points. Therefore, we can conclude that 0.147 increase in probability of COVID was mediated through chills. Practically, this says that 0.147/0.288 =51% of the effect of fever on COVID-19 is mediated through chills.

**Table 6. Mediated effect of fever on COVID-19 through Chills**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strata | Strata for Removing Confounding | Counterfactual Network without Fever🡪 Chills | Network with Best Fit to Datawith Fever 🡪 Chills | Freq of strata  |
| Below AverageAge | Female | Shivering | Fatigue | Loss of Taste | Prob of COVID, if Fever | Prob of COVID, if no Fever | Change in Prob of COVID | Prob COVID, if Fever | Prob COVID, if no Fever | Change in Prob of COVID |
| 1 | Yes | Yes | Yes | Yes | Yes | 0.493 | 0.486 | 0.007 | 0.49 | 0.483 | 0.007 | 0.000009315 |
| 2 | Yes | Yes | Yes | Yes | No | 0.475 | 0.482 | -0.007 | 0.472 | 0.476 | -0.004 | 0.000301185 |
| 3 | Yes | Yes | Yes | No | Yes | 0.484 | 0.483 | 0.001 | 0.481 | 0.481 | 0 | 0.000083835 |
| 4 | Yes | Yes | Yes | No | No | 0.475 | 0.457 | 0.018 | 0.471 | 0.449 | 0.022 | 0.002710665 |
| 5 | Yes | Yes | No | Yes | Yes | 0.499 | 0.483 | 0.016 | 0.501 | 0.484 | 0.017 | 0.000922185 |
| 6 | Yes | Yes | No | Yes | No | 0.481 | 0.458 | 0.023 | 0.483 | 0.441 | 0.042 | 0.029817315 |
| 7 | Yes | Yes | No | No | Yes | 0.482 | 0.482 | 0 | 0.481 | 0.481 | 0 | 0.008299665 |
| 8 | Yes | Yes | No | No | No | 0.459 | 0.3 | 0.159 | 0.465 | 0.154 | 0.311 | 0.268355835 |
| 9 | Yes | No | Yes | Yes | Yes | 0.497 | 0.496 | 0.001 | 0.49 | 0.49 | 0 | 0.000004185 |
| 10 | Yes | No | Yes | Yes | No | 0.494 | 0.494 | 0 | 0.49 | 0.488 | 0.002 | 0.000135315 |
| 11 | Yes | No | Yes | No | Yes | 0.494 | 0.494 | 0 | 0.49 | 0.49 | 0 | 0.000037665 |
| 12 | Yes | No | Yes | No | No | 0.492 | 0.466 | 0.026 | 0.489 | 0.456 | 0.033 | 0.001217835 |
| 13 | Yes | No | No | Yes | Yes | 0.497 | 0.496 | 0.001 | 0.494 | 0.491 | 0.003 | 0.000414315 |
| 14 | Yes | No | No | Yes | No | 0.49 | 0.484 | 0.006 | 0.491 | 0.478 | 0.013 | 0.013396185 |
| 15 | Yes | No | No | No | Yes | 0.491 | 0.491 | 0 | 0.49 | 0.49 | 0 | 0.003728835 |
| 16 | Yes | No | No | No | No | 0.487 | 0.35 | 0.137 | 0.487 | 0.201 | 0.286 | 0.120565665 |
| 17 | No | Yes | Yes | Yes | Yes | 0.494 | 0.493 | 0.001 | 0.485 | 0.484 | 0.001 | 0.000011385 |
| 18 | No | Yes | Yes | Yes | No | 0.487 | 0.483 | 0.004 | 0.477 | 0.472 | 0.005 | 0.000368115 |
| 19 | No | Yes | Yes | No | Yes | 0.489 | 0.489 | 0 | 0.484 | 0.484 | 0 | 0.000102465 |
| 20 | No | Yes | Yes | No | No | 0.475 | 0.457 | 0.018 | 0.469 | 0.446 | 0.023 | 0.003313035 |
| 21 | No | Yes | No | Yes | Yes | 0.491 | 0.495 | -0.004 | 0.485 | 0.49 | -0.005 | 0.001127115 |
| 22 | No | Yes | No | Yes | No | 0.469 | 0.445 | 0.024 | 0.46 | 0.434 | 0.026 | 0.036443385 |
| 23 | No | Yes | No | No | Yes | 0.486 | 0.486 | 0 | 0.484 | 0.484 | 0 | 0.010144035 |
| 24 | No | Yes | No | No | No | 0.47 | 0.297 | 0.173 | 0.469 | 0.115 | 0.354 | 0.327990465 |
| 25 | No | No | Yes | Yes | Yes | 0.497 | 0.497 | 0 | 0.493 | 0.493 | 0 | 0.000005115 |
| 26 | No | No | Yes | Yes | No | 0.492 | 0.492 | 0 | 0.485 | 0.484 | 0.001 | 0.000165385 |
| 27 | No | No | Yes | No | Yes | 0.495 | 0.495 | 0 | 0.493 | 0.493 | 0 | 0.000046035 |
| 28 | No | No | Yes | No | No | 0.489 | 0.463 | 0.026 | 0.485 | 0.45 | 0.035 | 0.001488465 |
| 29 | No | No | No | Yes | Yes | 0.498 | 0.495 | 0.003 | 0.497 | 0.493 | 0.004 | 0.000506385 |
| 30 | No | No | No | Yes | No | 0.484 | 0.48 | 0.004 | 0.493 | 0.493 | 0 | 0.016373115 |
| 31 | No | No | No | No | Yes | 0.494 | 0.494 | 0 | 0.493 | 0.493 | 0 | 0.004557465 |
| 32 | No | No | No | No | No | 0.468 | 0.313 | 0.155 | 0.468 | 0.121 | 0.347 | 0.147358035 |
| Freq | 0.45 | 0.69 | 0.01 | 0.10 | 0.03 | Change in COVID =  | 0.141 | Change in COVID = | 0.288 | 1 |