Sepsis Distribution in the US States

Raehash Shah

2023-11-17

Introduction

Sepsis is a life-threatening condition that affects 1.7 million adults in the United States every year. Sepsis is typically a result of an infection, which are commonly bacterial, triggering a cascade of reactions throughout the body. These infections, which often begin in the lungs, urinary tract, skin or the gastrointestinal tract, can cause tissue damage, organ failure or death if not properly treated (CDC, 2023). Given that so many Americans are affected by Sepsis, we would like to observe what the distribution of Sepsis is across the U.S. states in 2023.

Methodology

As previously mentioned, bacterial infections are the most common cause of Sepsis. This is further supported, as the three germs that most frequently develop into Sepsis are Staphlyococcus aureus, Escherichia coli (E. coli) and some type of Streptococcus (Johns Hopkins Medicine, 2019). Thus, three data sets recorded by the National Notifiable Disease Surveillance System (NNDSS) in 2019 on the number of Vancomycin-resistant Staphylococcus aureus, Shiga toxin-producing E. coli and Streptococcal toxin shock syndrome infections across the 50 states (NNDSS, 2019) gave an estimate on the number of each bacterial infection in each state which was later used to create an estimate on number of patients with Sepsis.

Although each of these data sets don't record all of the bacterial infections caused by each of these bacteria individually in each state, these data sets still cover the relevant bacterial infections that cause sepsis: (1) Vancomycin is an antibiotic that treats bacterial infections which means resistant bacterial infections were resistant to the antibiotic and thus could have developed into Sepsis (Cong et.al, 2020), (2) Shiga toxin-producing E. coli are bacteria that have been shown to be a variant of E. coli that leads to Sepsis (Braune et.al, 2013), and (3) Streptococcal toxin shock syndrome infections can often also cause Septic shock which is a result of Sepsis which makes this data relevant to our estimate.

However, not all bacterial infections cause Sepsis. In fact from a study performed in 2002, roughly 53% of the sepsis patients were caused by gram positive or gram negative bacteria (Alberti et.al, 2001). In addition, of those Sepsis cases caused by bacterial infections, roughly 62% are a result of gram positive bacteria and 47% are a result of gram negative bacteria (Dolin, 2019). Since Staphylococcus aureus and Streptococcus are gram positive bacteria and E. Coli is gram negative, the number of patients diagnosed with each bacteria in each state was normalized based on the class of the bacteria so that the final number of each bacteria infections in each state are representative of the number of Sepsis cases that are predicted based on if Sepsis is caused by bacteria and the class of the bacteria.

In addition to an approach done using the bacteria that cause Sepsis, Sepsis has also been shown to be correlated with certain diseases. In fact most people that develop Sepsis have at least one underlying condition (CDC, 2023). Among the common diseases associated with Sepsis include Chronic Pulmonary Disease, Chronic Kidney Disease, Coronary Heart Disease, Diabetes, Hypertension and Stroke (Wang et.al, 2012). Therefore, data taken on number of patients with each diagnosis in 2017 by the 500 cities project (500 cities, 2017) were used to calculate the number of patients diagnosed with each of these diseases that year.

Since not all patients with these diseases end up with Sepsis, the number of patients estimated is scaled by the risk of sepsis given that a patient has the disease which was published in Wang et.al, 2012. From there the number of patients is normalized so that the total number of patients equal the total number of patients estimated to be diagnosed with Sepsis each year.

For both data frames, the data frames were created using the python script "setup_dataframes.py" which used the techniques described above to create a final data frame that had each state and a value that represented an estimate on the number of people in the state affected by Sepsis. Since this is an estimate, the value for each state wasn't rounded since it is not an exact number of people with Sepsis per state. Then the data frame was loaded directly into R/Rmarkdown to create a final report and plot the distribution by state.

Results

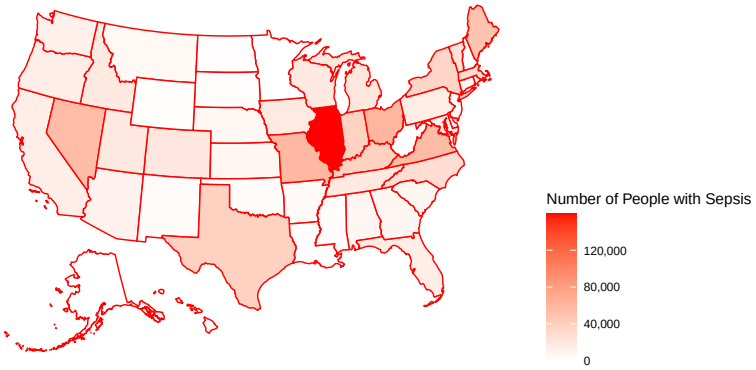

Figure 1: Distribution of Sepsis caused by Bacterial Infections in U.S. States

The above heat map is a result of plotting estimates in the number of Sepsis patients based on bacterial infections. From our results we observe that there isn't a distinct pattern regionally that separate states with Sepsis from states that don't. However, it does seem that states with cities are more likely to have patients with Sepsis (ex: Chicago in Illinois, Las Vegas in Nevada, etc.). More exploration is necessary in order to confirm that hypothesis.

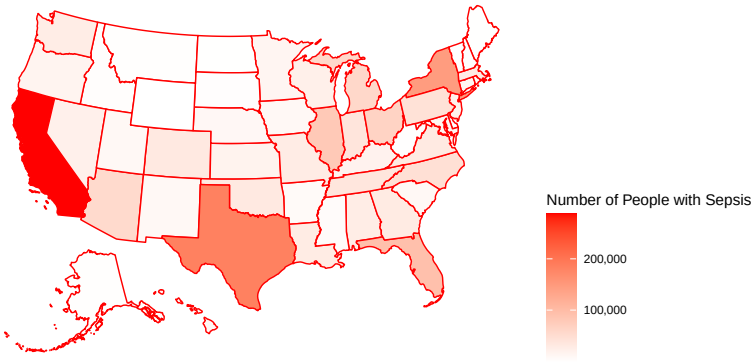

Figure 2: Distribution of Sepsis Based on Correlated Disease Prevalence in U.S. States

In this above heat map we estimate Sepsis prevalence in a state with the number of patients diagnosed with a correlated disease. Note that each state's correlated disease prevalence was scaled so that the total number of patients with Sepsis in this heat map add up to the 1.7 million annual Sepsis patients reported by the CDC. Like our distribution of Sepsis based on number of bacterial infections, this heat map doesn't seem to follow a regional based pattern but rather based on cities. However, it is possible since the data for this heat

map was performed by the 500 cities project that we simply have more data regarding patients in cities than those patients who are in more rural regions of U.S. which may indicate that we can't make that conclusion for this heat map. Comparing the two estimates we created we have some clear differences - in this heat map we observe California to be the state estimated to have the most number of people with Sepsis where on the first map the estimate was Illinois. On the other side, there are some consistencies in relative prevalence of Sepsis in states we see between the heat maps. States that have a high or low prevalence in both can be indications that we are confident about their relative prevalence since using two different metrics the results are the same.

Limitations

Our estimates that were created for the number of Sepsis patients came with some assumptions. However, these assumptions can limit how we interpret the estimates. For our distribution estimated in Figure 1, a limitation is that this doesn't represent all Sepsis patients in the U.S.: rather just those caused by bacterial infections. Therefore, we can't assume with certainty that the distribution of Sepsis follows the distribution of Sepsis from bacterial infections. In addition, our estimates on the number of Sepsis patients from bacterial infections only used the top three bacterial infections. In reality, research suggests there are many more so with more data and incorporating that data into our model would provide a better estimate on the number of Sepsis patients from bacterial infections. Similarly in Figure 2, the limitation is the data set used. Only a subset of all diseases correlated with Sepsis were used to estimate the total number of sepsis patients and more data will improve our estimate model. Finally, for both models, it is important to consider the date of our data sets. 2017 and 2019 are both pre-COVID, and it has been shown that COVID can lead to Sepsis. This suggests that as we live in a post-COVID world, there may be a different distribution. Hence, provided with more updated data the estimates will be more representative of the true number of patients with Sepsis in each state in 2023.

References

500 Cities: Local Data for Better Health, 2017 Release. Publish on Centers for Disease Control & Prevention. (https://data.cdc.gov/500-Cities-Places/500-Cities-Chronic-obstructive-pulmonary-disease-a/xzgm-d4ya, https://data.cdc.gov/500-Cities-Places/500-Cities-Chronic-kidney-disease-among-adults-age/dnkc-3whb, https://data.cdc.gov/500-Cities-Places/500-Cities-Stroke-among-adults-aged-18-years/se7n-wpp3, https://data.cdc.gov/500-Cities-Places/500-Cities-Diagnosed-diabetes-among-adults-aged-18/cn78-b9bj, https://data.cdc.gov/500-Cities-Places/500-Cities-Coronary-heart-disease-among-adults-age/cqcq-r6f8, https://data.cdc.gov/500-Cities-Places/500-Cities-High-blood-pressure-among-adults-aged-1/ebxs-yc6e)

Alberti, Corinne, et al. "Epidemiology of sepsis and infection in ICU patients from an international multicentre cohort study." Intensive Care Medicine, vol. 28, no. 2, 2001, pp. 108–121, https://doi.org/10.1007/s00134-001-1143-z.

Braune, Stephan A., et al. "Clinical features of critically ill patients with shiga toxin–induced hemolytic uremic syndrome." Critical Care Medicine, vol. 41, no. 7, 2013, pp. 1702–1710, https://doi.org/10.1097/ccm.0b013e31828a24a8.

Cong, Yanguang, et al. "Vancomycin resistant Staphylococcus aureus infections: A review of case updating and clinical features." Journal of Advanced Research, vol. 21, 2020, pp. 169–176, https://doi.org/10.1016/j.jare.2019.10.005.

Dolin, Hallie H, et al. "Characterization of pathogenic sepsis etiologies and patient profiles: A novel approach to triage and treatment." Microbiology Insights, vol. 12, 2019, p. 117863611882508, https://doi.org/10.1177/1178636118825081.

National Notifiable Diseases Surveillance System (NNDSS), 2019 Release. Published by Division of Health Informatics and Surveillance in the Centers for Disease Control and Prevention. (https://data.cdc.gov/NNDSS/NNDSS-TABLE-1KK-Vancomycin-resistant-Staphylococcu/ybum-psnc,

https://data.cdc.gov/NNDSS/NNDSS-TABLE-1FF-Severe-acute-respiratory-syndrome-/5wqm-pm27, https://data.cdc.gov/NNDSS/NNDSS-TABLE-1HH-Streptococcal-toxic-shock-syndrome/ap9g-4wiq)

"Septicemia." Johns Hopkins Medicine, 19 Nov. 2019, www.hopkinsmedicine.org/health/conditions-and-diseases/septicemia.

Wang, Henry E., et al. "Chronic medical conditions and risk of sepsis." PLoS ONE, vol. 7, no. 10, 2012, https://doi.org/10.1371/journal.pone.0048307.

"What Is Sepsis?" Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 24 Aug. 2023, www.cdc.gov/sepsis/what-is-sepsis.html.