Medical Staff Allocation in 2018

Interim report

The United States has an influenza season where more people than usual suffer from the flu. This report investigates patterns and correlations to decide how to allocate temporary medical staff across the states for the upcoming season.

by E.Wojnicka

Project Overview

Motivation

The United States has an influenza season where more people than usual suffer from the flu. Some people, particularly those in vulnerable populations, develop serious complications and end up in the hospital. Hospitals and clinics need additional staff to adequately treat these extra patients. The medical staffing agency provides this temporary staff.

Objective

Determine when to send staff, and how many, to each state.

Scope

The agency covers all hospitals in each of the 50 states of the United States, and the project will plan for the upcoming influenza season.

Research hypothesis

According to the "Demographic Turning Points for the United States: Population Projections for 2020 to 2060", American society is aging rapidly, and the fraction of the age group 65+ is expanding. For this reason, the rest of this research will focus on the population of people 65 and older. First, the goal is to establish when and where to send the additional medical staff to speed up vaccination and cover the vaccination reach more effectively. Secondly, we would like to know which hospitals need our support during the influenza outbreak.

Hypothesis 1.

Highly populated states with a population containing larger fractions of the age group 65+ will need more staff allocation because of the low medical staff vs. population ratio.

Hypothesis 2.

States with low vaccination rates and high deaths rates

will need more staff allocation before the influenza season to prevent the increase in death cases.

Hypothesis 3.

Some states have high population, low ratio of medical staff vs population, low ratio of hospitals vs population and low rates of vaccination which result in high death rates within total population.

Data overview

<table>
<tr><td colspan="2">Data set "Influenza deaths by geography, time, and age"</td></tr>
<tr><td>Data source</td><td>Center for Disease Control and Prevention, National Center for Health Statistics</td></tr>
<tr><td>Data collection method</td><td>Based on the death certificates records for U.S. citizens. Each death certificate identifies the death cause with help of ICD-10 codes.</td></tr>
<tr><td>Data overview</td><td>Database contains mortality rates by state. Provided dataset reflects data collected between 2010 and 2019.</td></tr>
<tr><td>Categories</td><td>USA states (excluding Puerto Rico), time (months and years), 5-year age groups and count of death cases reported.</td></tr>
<tr><td><strong style="color:red">Limitations</td><td>Data doesn't include "Total influenza patients", "Ethnic" and "Risk diseases" categories, which is relevant to get a very accurate insights on influenza deaths distribution.</td></tr>
<tr><td colspan="2">Data set "Population data by geography"</td></tr>
<tr><td>Data source</td><td>US Census Burau</td></tr>
<tr><td>Data collection method</td><td>US Census Bureau measures population change (birth, deaths, domestic/international migration) and adds it to the base population to update estimates every year. The in-depth description of the   methodology for the USA population estimates is available in the separate publication</td></tr>
<tr><td>Data overview</td><td>Data set is an estimate in years 2009-2017.</td></tr>
<tr><td>Categories</td><td>Total population, population by sex differentiation, 5-years age groups.</td></tr>
<tr><td><strong style="color:red">Limitations</td><td>There where duplicates in State and Year categories but different values in Population estimates. These duplicates were removed automatically with Excel function. The numbers, however, doesn't represent the real value. Data set doesn't reflect ethnic structure of each state, which excludes the research are by race.</td></tr>
</table>

Data set "Flu Surveillance"

Data source	Center for Disease Control and Prevention
Data collection method	FluSurv-NET conducts population-based surveillance for laboratory-confirmed influenza-associated hospitalizations in children (persons younger than 18 years) and adults. The current network covers over 70 counties in the 10 Emerging Infections Program (EIP) states (CA, CO, CT, GA, MD, MN, NM, NY, OR, and TN) and four additional states (IA, MI, OH, and UT) through the Influenza Hospitalization Surveillance Project (IHSP). Approximately 9% of US population is covered by this surveillance system.
Data overview	Case counts and rates for hospital admissions due to influenza infections. The data base also includes information about vaccination.
Categories	Geography, ages groups, types of vaccination, season
Limitations	Data bank difficult to filter and retrieve relevant information. The data set has low coverage.

Data set "Hospitals in USA"

Data source	Homeland Infrastructure Foundation-Level Data (HIFLD)
Data collection method	Data acquired from various state departments or federal sources which has been referenced in the Source field.
Data overview	This feature class/shapefile contains locations of Hospitals for 50 US states, Washington D.C., US territories of Puerto Rico, Guam, American Samoa, Northern Mariana Islands, Palau, and Virgin Islands.
Categories	Hospitals have been categorized into children, chronic disease, critical access, general acute care, long term care, military, psychiatric, rehabilitation, special, and women. Other categories include state code, county, address, phone number and number of beds.

Limitations	The database does not contain nursing homes or health centers.

Number of hospitals and hospital employment in each state in 2019

Data source	U.S. Bureau of Labor Statistics
Data collection method	The Current Employment Statistics-National (CES-N) program produces detailed industry estimates of nonfarm employment, hours, and earnings of workers on payrolls. Each month, BLS collects data through the Current Employment Statistics (CES) survey, which includes about 131,000 businesses and government agencies, representing approximately 670,000 individual worksites drawn from a sampling frame of unemployment insurance tax accounts covering roughly 10.4 million establishments.
Data overview	Compares the number of hospitals with the size of the population and the number of hospital employees with the size of the population.
Categories	Sates, Ratio of people per hospital, Hospital establishments in third quarter 2019, Population (2019), Ratio of people per hospital employee, Hospital employment 2019.
Limitations	Data set considers only one year, that is not included in primary data sets.

Descriptive Statistics Summary

Core variables overview

Vaccination (season 2016-2017) – <u>no statistics, data set need a second review</u>

Variable		Min	Max	Average	Standard deviation
Vaccinated total		n/a	n/a	n/a	n/a

| Vaccinated 65+ | | n/a | n/a | n/a | n/a |

Population (2017)

Variable	Min	Max	Median	Standard deviation
Under 5	15618	274924	169671	382706
5-64	209490	31002572	3128816	4640236
65+	31577	527352	527352	737746

Influenza (2017)

Variable	Min	Max	Average	Standard deviation
Patients total	37129	5158325	766340	1017530
Deaths 65+	216	5694	942	955
Deaths 5-65	469	1307	621	108

Medical Care (2018-2020)

Variable	Min	Max	Average	Standard deviation
Hospitals establishments	24	1555	272	283
Medical staff	12332	582706	127220	126645

Correlations

Variables	Correlation value	Description
Age group 65+ and death rates	0,872 — strong	Age groups 65+ are exposed to the high risk of dying on influenza
Hospitals ratio to population and influenza death rates	0.842 — strong	States with low hospitals ratio to the population correlate negatively on deaths rates

Results summary and insights

<table>
<tr><td colspan="2" style="color:green">Research hypothesis</td></tr>
<tr><td colspan="2">The bigger fraction of 65+ years old people in a state results in proportionally bigger numbers of death cause by influenza</td></tr>
<tr><td>Depended variable</td><td><code>Age group</code></td></tr>
<tr><td>Independent variable</td><td><code>Influenza deaths</code></td></tr>
<tr><td>Null hypothesis</td><td><code>Deaths caused by influenza for the group 65+ years old is <= than for the groups 5-64 years old</code></td></tr>
<tr><td>Alternative hypothesis</td><td><code>Deaths caused by influenza for the group 65+ years old is > than for the groups 5-64 years old</code></td></tr>
<tr><td>Type</td><td><code>One tail test</code></td></tr>
<tr><td>Alpha</td><td align="right"><code>0.05</code></td></tr>
<tr><td>p-value</td><td align="right"><code>0,024</code></td></tr>
<tr><td>Significance level</td><td><code>The p-value is less than alpha. The null hypothesis can be rejected. Influenza deaths rates for population in age 65+ is grater than the population 5-64 years old wiuth 95% confidence</code></td></tr>
<tr><td>Test results</td><td>
<pre>
t-Test: Two-Sample Assuming Unequal Variances

                                    748   3689352,7

Mean   1877,28105   6320028,21

Variance   404236274   4,6729E+15

Observations   459   459

Hypothesized Mean Difference   0

df   458

t Stat   -1,9801816

P(T<=t) one-tail   0,02414058

t Critical one-tail   1,64818742

P(T<=t) two-tail   0,04828116

t Critical two-tail   1,9651571
</pre>
</td></tr>
<tr><td>Conclusion</td><td style="background:green;color:white"><code>Hypothesis confirmed.</code></td></tr>
</table>

Remaining analysis

Further investigations

1. Which states have the highest fraction of the population from the age group 65+
2. Correlations in the coverage of vaccination in 2016
3. Hospitals to population ratio
4. Medical staff to population ratio
5. Influenza outbreaks peak time
6. Types of places where people get vaccinated

Visualizations to consider

1. Distribution of 65+ years old group by state
2. Influenza outbreak at the beginning of the season, in the peak point and at the end
3. Population ratio to the number of private and government hospital establishments, by state
4. Population ratio to the number employed in private and government hospitals.
5. Frequency of deaths caused by influenza by age group (5-64 and 65+)
6. Vaccination distribution in risk or no-risk groups, by state