Analysis of Gender Differences in COVID-19

Izabela Mitreska[1], Oliver Simonoski[1], Rasim Salkoski[1]

[1]The University of Information Science and Technology "St. Paul the Apostle", Ohrid, North Macedonia
{izabela.mitreska@cse.uist.edu.mk,
oliver.simonoski@cse.uist.edu.mk, rasim.salkoski@uist.edu.mk}

Abstract. The Covid-19 pandemic epidemiological data should be provided by age and sex segments, according to the global health community, but this specification which is degraded by age and sex hardly reported. There is high importance in such information's but also, they are essential for the population to make properly informed decisions regarding their personal illness risk, as well as for governments to construct public policy. Our paper aims to investigate the relationship and influence of the experimental design between the gender-specific Covid-19 infections and losses. To create short-term projections of the Covid-19 outbreak over time period, we set the following hypothesis: "As the spread of Covid-19 increased, the number of infected women also increases". We have used R Studio as an integrated development environment for R to verify the accuracy of the hypothesis. The layout of the publication analyzes related ideas regarding the variations and effects of Covid-19 infections based on sex and gender. In order to estimate the evaluation of represented experimental design we used a dataset to demonstrate that women experience higher infection rates than male population. Determined through the calculation of essential statistical values we drawn conclusion that the male population had a higher death rate than the female population.

Keywords: Covid-19, R Studio, dataset, experimental design, hypothesis

1 Introduction

The key patient health data for this research, the Coronavirus disease 2019 (Covid-19), is an infectious disease caused by a recently discovered virus that was spread among humans, other countries and among different cities, mainly in China, since December 2019 [1].

There were significant gender differences in testing and case rates during the early phases of the Covid-19, these variations are not connected only to the age differences, but also to the general state of the subjects' health [2].

Researchers in the paper [3] has proven that there are an equal number of cases between sexes who are infected with Covid-19. The current evidence indicates that there are higher infection severity and mortality in men rather than women [4].

The aforementioned claims from different investigators are the major reasons that lead us to create an experimental design which will demonstrate our hypothesis: "*As the spread of Covid-19 increased, the number of infected women also increased*", which is contrary to other claims. In order to achieve success results in this process, to prove or reject our hypothesis, we will use a specific data collection represented in Section 4.

The paper is organized as follows. Section 2 represents different concepts related on the newly detected Covid-19 in both, female and male, genders. Using some experimental skills, such as classifying, inferring and predicting, we produce the experimental design represented in Section 3. The dataset and tools which are needed for the process are represented in Section 4, followed by Section 5 describing the healthcare patient data in both, female and male, genders. The Information about Covid-19 protection measures are provided in Section 6 and Section 7 represents the final conclusions.

2 Related Work

Many researches continue to demonstrate the differences in severity and effects of Covid-19 infections. Hospitalized patients, regardless of the gender, are less likely to die from the novel virus. However, further studies are needed to better investigate the outcomes of Covid-19 [5].

Although available sex-disaggregated data for Covid-19 show equal numbers of cases between sexes, current evidence indicates that fatality rates and certain comorbidities are associated with higher infection severity and mortality in men rather than women [6, 7]. These findings provide a possible explanation for the observed sex biases in Covid-19 and an important basis for the development of a sex-based approach to the treatment of male and female patients [8].

The outcomes of this procedures [9, 10] represents that women, worried more about their loved ones and their health, are more likely to see Covid-19 as a very serious health problem. On the other hand, the male population were more occupied with effects on the economy and society and they were infected less than the female gender [11].

However, the government respond with rapid and severe steps to stem the pandemic's spread [12]. This is critical to recognize how these policies may disproportionately increase the risks for women, both directly and indirectly, in terms of gender [13].

Unfortunately, most research papers consistently fail to provide comparisons between men and women, but still provide that there are lower vaccination intentions among different genders [14, 15], which is contrary to the conclusion indicated in the paper [16].

The process represented in this paper aims to quantify the Covid-19 pandemic's dynamic parameters, including the need for precise and timely data collecting that distinguishes and categorizes Covid-19 affected individuals. The represented outcomes in this paper can help to confirm and define the differences in infection and fatality rates. This data is important for stakeholders at the national and subnational levels who are responsible for developing and implementing a Covid-19 exit strategy.

3 Experimental Design

In order to get precise, maximized and clear data for further research we produced an experimental design, a scientific approach to data collection and measurement, from where we can derive our hypothesis: "*As the spread of Covid-19 increased, the number of infected women also increases*". In fact, our overall goal is to prove that fatality rates and certain comorbidities from Covid-19 are associated with higher infection in women rather than men.

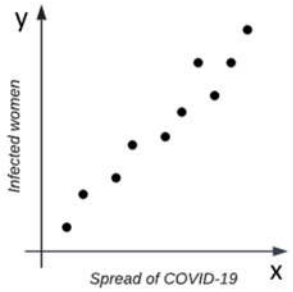

Figure 1: As the spread of COVID-19 increased, the number of infected women also increases.

There are a lot of concepts and terms related to the experimental design, but in our procedure we will put the main emphasis only on the independent and dependent variables.

The independent variable, spread of Covid-19 in our case, it is the variable that we continuously manipulate and displayed on the x-axis because does not depend on other variables which are being measured.

On the y-axis, we determined the dependent variable that adjust to the changes in the independent variable. This process means that changes in x, the independent variable, effect changes in y (see Fig. 1).

Hence, from the figure above (see Fig. 1) we can see that as the virus spread increases, the number of infected woman also increased which means that changes in the independent variable, spread of Covid-19, effect changes on the dependent variable, the increased number of infected women.

If we want to maintain control and keep the data consistently updated it is needed to follow the instant spread of the Covid-19 pandemic. The dataset used in our process confirms a clear relationship between settlement population gender and viral infections across populations.

This Section provides the ability to discard unnecessary data, secure and improve the performance which is a crucial stage in the data collection process in order to be prepare for further researches in the following Section.

4 Experimental Platform

4.1 Software

The development and implementation of our process are derived with the support of R Studio, a graphical user interface for R. R is a powerful programming language and environment used in statistical computation, data analytics, scientific research, and graphics which offers built-in functions that makes the application a great tool for debugging and workspace management.

4.2 Dataset

In our process we used the "COVID19_line_list_data" dataset which was downloaded from the Kaggle website [17].

Table 1: Specific aims and function of researching

Specific aim	The function of researching
Table representation of the dataset and the main hypothesis	*Representing the correlation between the confirmed cases and the proposed hypothesis*
Graph representation of the correlation between male and female population - alive and death cases	*Total cases - male vs. female*

Box Plot representation of gender ratio	*Calculating and processing the difference between the recovered and death cases*

Table 1 represents the main aim and impact through the whole process. Through the aforementioned software and dataset, we will see in the further section if the proposed hypothesis will be confirmed or rejected. Choosing a good software for analysis plays a key role in the whole process for which a good preliminary research is needed.

5 Data Processing

The dataset used in our procedure contains daily data on the number of impacted cases, deaths and Covid-19 patient recovery. Our outcomes and conclusion may be outdated since Covid-19 occurrences and rates are constantly changing.

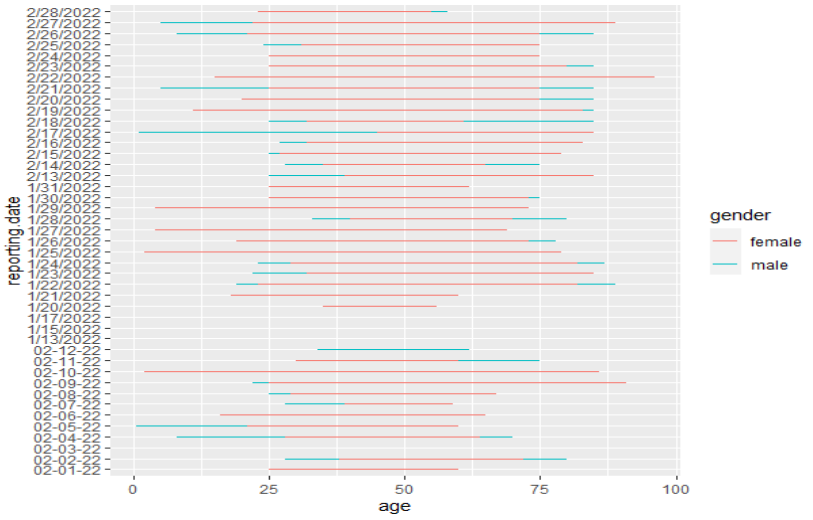

Figure 2: Representation of impact of Covid-19 in different ages

Through the implementation of the software and tools mention in Section 4, we performed a data analysis which outcomes will show how the virus spread among humans depending on their age. On the x-axis, we determined the age of people who are affected from Covid-19 in a range from 0 to 100 age and on the y-axis we represented the date when people were infected from this virus. The results which we obtained from our process show that female are more affected from 30 to 50 age and male are more affected in a range from 0 to 30 and from 50 to 100 age (see Figure 2).

```
------------------------------------------------
gender
       n  missing distinct
     902      183        2

Value      female   male
Frequency     382    520
Proportion  0.424  0.576
------------------------------------------------
```

Figure 3: Displaying the frequency and proportion of the epidemic by gender

In the primary visualization we represent the structure of the epidemic by gender and provide different information about the number of samples, missing and distinct values and their frequency and proportions. As a first step, we split the given dataset by gender and calculate all of the attributes represented on Figure 3. The next step was to calculate the mean which shows that the male population had a greater death rate (8.5%), with 520 frequency and 0.576 proportion, than the female population (3.7%), with 382 frequency and 0.424 proportion. We can say that this was unexpected, which leads us to the question: "*Can we trust this info*?".

```
        Welch Two Sample t-test

data:  men$death_dummy and women$death_dummy
t = 3.084, df = 894.06, p-value = 0.002105
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.01744083 0.07849151
sample estimates:
 mean of x  mean of y
0.08461538 0.03664921
```

Figure 4: Implementation of R environment's built-in function

To verify the accuracy of the obtained results, we use the R environment's built-in function. Our 95 percent confidence interval indicates that male deaths are at a rate ranging from 1.7 percent to 7.8 percent greater than female. The null hypothesis that men and women had the same death rate may be rejected since the p-value of 0.05 is more than 0.0002 (See Figure 4).

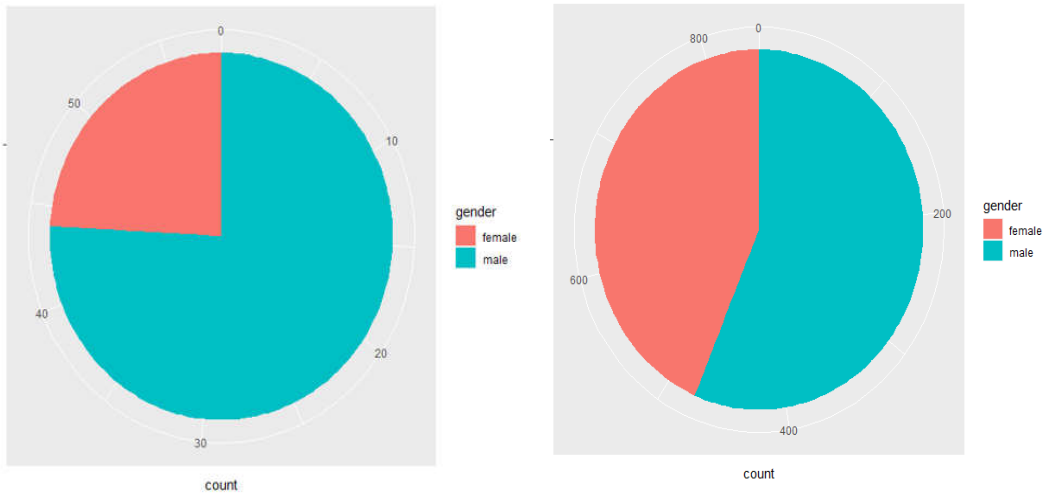

Figure 5: Graphical figure depicting the relationship between the population's male and female genders in perspective including both alive and dead cases

The findings of multiple sensitivity analyses conducted using different classifications based on absolute measures of settlement population size are depicted in the graphs above. Specifically, we represented the relation between the total number of cases divided by gender (alive and death cases).

As we can see from the charts above (see Figure 5), females have a slightly smaller percentage of cases compared to the male population. From the total number of the confirmed cases, around 480 cases from the recovered belong to the male gender, while the mortality rate in the female population is lower compared to the male population.

The box plot represented on the figure bellow (see Figure 6) show that the number of deaths compared to the recovered one is much smaller, which is considered a quality result. The black dots represented on the 1.0 y-axis give the main pictures about the current situation which is similar to the previous outcomes, but in this process the visual ties are much better and understandable. The first and the third quartile for both female and male population is almost the same without some major concessions.

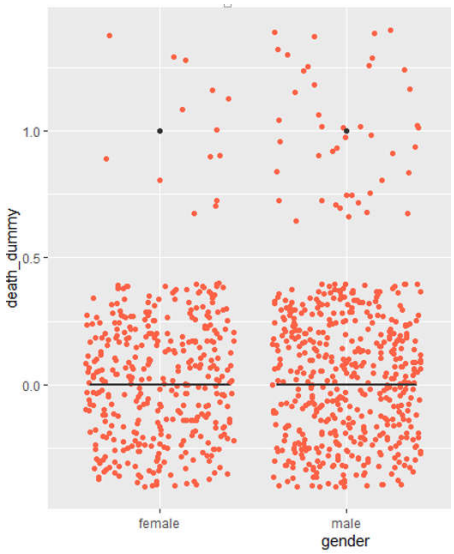

Figure 6: Box Plot representation of gender ratio

Our findings imply that Covid-19 infection and fatality rates are divided by age and gender, which does not necessarily represent population structure. From the beginning of the pandemic, it was clear that Covid-19 severity increased with age. One possible explanation is because women represent a large percentage of workers in the health-care industry and care institutions, and hence are at the forefront of the Covid-19 struggle. Furthermore, working in health and care facilities increases their chances of being tested. The observed gender differences highlight the importance of paying additional attention to health-care employees, regardless of gender.

However, we gained the information that most of the reported cases belong to the male population, but their age limit is quite close with small variations in both, male and female populations which is contrary to our hypothesis and typically for further places in the world.

6 COVID-19 Protection Measures

After a crisis, recovery periods might provide an opportunity to do things differently, try new things and create more justified systems and civilizations. Understanding gender inequalities and preventing backsliding will be necessary for achieving equality and preventing backsliding. Society has progressed as a response to the crisis, the produced inequitable gender standards and the effects of the pre-infection. However, the main goal is to conduct gender impact evaluations that identify strengths and problems.

While men had higher death rates, the economic and social consequences have amplified gender inequities and disproportionately harmed women's employment security: "Women form the bulk of vulnerable employees who are losing their jobs and salaries at an alarming rate".

In order to guarantee the response preparations for COVID-19 and any future pandemics or global emergencies, we propose the following guidelines:

1. National governments should ensure that response strategies include the following elements:
    a. Disaggregated data on gender;
    b. Incorporating gender considerations and applying gender analysis;
2. Make safety a top priority;
3. Challenge inequitable work distribution and restrictive gender roles;
4. Ensure that everyone has access to health care;
5. Gender equality should be a priority in monitoring and assessment;

Table 2: Re-establishing a better basis for expanding social protection systems requires

Implement shock-responsible social protection systems that prioritize rights and dignity
Support the creation of strong care systems that share the expenses of care with society as a shared obligation
Ensure active participation of women in decision-making, including through meaningful collaboration with local organizations
Invest in data collection for inclusive and gender-responsive policies across all social protection areas

7 Conclusion

The development of experimental design has made the work of the data scientists easier in monitoring health related data of individuals suffering from various diseases. In order to demonstrate that greater infection rates are in women rather than men which are associated with higher mortality rates and specific comorbidities from Covid-19, we developed this methodology for accurate and fast data gathering.

We analyze Covid-19 cases that have been presented across different genders in different places using suitable dataset, tools and objectives. Through the outcome of our process we gained the information that there were far more male patient cases, than female patient cases, which is contrary to our proposed hypothesis.

References

1. Sultana, S. at el.: Gender differences in knowledge, attitudes and preparedness to respond to Covid-19 among adults in Bangladesh: A cross-sectional study. Population Medicine, vol. 4, 12 January (2022)
2. Jewel, G., Langer A.: Sex and Gender Disparities in the Covid-19 Pandemic. Journal of Women's Health. vol. 29, no. 4, (2020)
3. Ancochea, J.: Evidence of Gender Differences in the Diagnosis and Management of Coronavirus Disease 2019 Patients: An Analysis of Electronic Records Using Natural Language Processing and Machine Learning. Journal of Women's Health, vol. 30, no. 3, (2021)
4. Spagnolo, P. at el.: Sex and Gender Differences in Health: What the Covid-19 Pandemic Can Teach Us. Annals of Internal Medicine, 1 September (2020)
5. Ya'qoub, L. at el.: Sex and gender differences in Covid-19: More to be learned", American Heart Journal Plus: Cardiology Research and Practice, 6 April (2021)
6. Raimondi, F. at el.: Covid-19 and gender: lower rate but same mortality of severe disease in women-an observational study. BMC Pulmonary Medicine (2021)
7. Jian-Min, J. at el.: Gender Differences in Patients with COVID-19: Focus on Severity and Mortality. Frontiers in Public Health. April (2020)
8. Bwire, G.: Coronavirus: Why Men are More Vulnerable to COVID-19 than Women. SN Comprehensive Clinical Medicine, pp. 874-876, 4 June (2020)
9. Takahashi, T. at el.: Sex differences in immune responses that underline COVID-19 disease outcomes. Nature, vol. 588, 26 August (2020)
10. Vegt, I. and Kleinberg, B.: Women Worry About Family, Men about the Economy: Gender Differences in Emotional Responses to COVID-19. International Conference on Social Informatics, vol. 12467, 7 October (2020)
11. Etheridge, B. at el.: The Gender Gap in Mental Well-Being during the Covid-19 Outbreak: Evidence from the UK. ISER Working Paper Series, no. 2020-08,7 June (2020)
12. Ambrosino, I. at el: Gender differences in patients with COVID-19: a narrative review. Monaldi Archives for Chest Disease 2020, vol. 90:1389 (2020)
13. Misra V. at el.: Gender disparity between authors in leading medical journals during the COVID-19 pandemic: a cross-sectional review. BMJ Open, 25 June (2021)
14. Zintel, S. at el.: Gender differences in the intention to get vaccinated against COVID-19: a systematic review and meta-analysis. Journal of Public Health, 7 January (2022)
15. Galasso, V. at el.: Gender differences in COVID-19 related attitudes and behavior: Panel evidence from eight countries. The National Bureau of Economic Research, 26 August (2020)
16. Dong, E. at el.: An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis, vol. 5, pp. 533-534 (2020)
17. Dong, E. at el.: An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis, vol. 5, pp. 533-534 (2020)