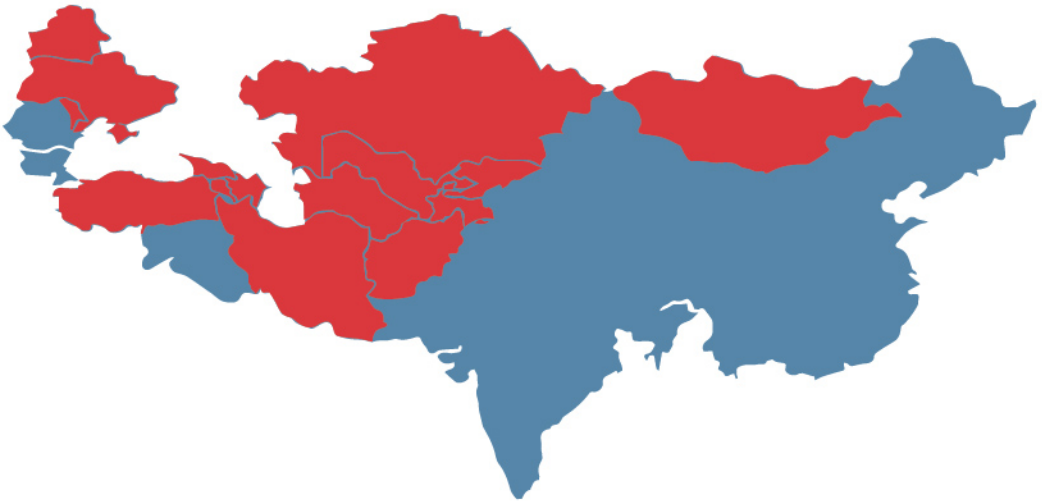
The Pan African Medical Journal

COVID-19 Outbreak in Herat Prison: An Exercise of Basic Biostatistical Analysis

Shoaib Naeemi, Khwaja Mir Islam Saeed, Mir Salamuddin Hakim

Learning from Practice: Public Health Teaching Case Studies from Eastern Europe and Central Asia



ISSN: 1937 – 8688

An Open Access Journal published by **The PAMJ**

EMPHNET
The Eastern Mediterranean Public Health Network

PanAfrican Medical Journal

COVID-19 Outbreak in Herat Prison: An Exercise of Basic Biostatistical Analysis

A Case Study

Authors: Shoaib Naeemi[1,&], Khwaja Mir Islam Saeed[2], Mir Salamuddin Hakim[3]

Affiliations: [1]Technical Officer, Eastern Mediterranean Public Health Network (EMPHNET) Afghanistan Country Office, [2]Technical Advisor for Afghanistan Field Epidemiology Training Program (AFETP), Eastern Mediterranean Public Health Network (EMPHNET) Afghanistan National Public Health Institute, Ministry of Public Health, [3]Technical Officer for Afghanistan Field Epidemiology Training Program (AFETP), Eastern Mediterranean Public Health Network (EMPHNET) Afghanistan National Public Health Institute, Ministry of Public Health

[&]Corresponding author: Shoaib Naeemi, MPH, Pharm-D, MLS

Email: snaeemi@emphnet.net

Phone-number: +93(0)764523311

Physical address: Cinema Pamir, Central Blood Bank, 5th Floor, Kabul City, Afghanistan

Abstract

Biostatistics plays a pivotal role in the field of epidemiology, particularly in the training programs aimed at equipping professionals with the necessary skills to effectively analyze and interpret data related to disease outbreaks. In the context of epidemiology, biostatistics provides a systematic framework for collecting, organizing, analyzing, and interpreting data, enabling epidemiologists to gain valuable insights into the patterns, determinants, and impact of diseases on populations. A comprehensive understanding of biostatistics is crucial for epidemiologists as they strive to identify risk factors, assess disease burden, and design appropriate public health interventions.

This practical exercise which highlights the application of biostatistics in epidemiology, is the analysis of a line list from an outbreak that occurred in Herat Province prison, Afghanistan in July 2022. Using the principles of biostatistics, participants would be tasked with conducting a basic analysis of this line list. The exercise would also emphasize the importance of data visualization techniques. Participants would be encouraged to create graphs, such as bar and pie charts, to visualize the temporal distribution of cases and identify potential patterns or clusters.

Through this exercise, participants would not only gain practical experience in utilizing biostatistical techniques but also understand the value of these methods in identifying key epidemiological features of an outbreak. By accurately analyzing the line list data, they would be able to identify risk factors, understand the mode of transmission, and inform public health interventions and control measures to mitigate the spread of the outbreak.

Time allotted: 1-2 Hours

Language: English

Groups: 3-5 residents per group

Materials Required: Pen, pencil, flip chart, calculator, marker, highlighter pen, ruler and personal computer (if available)

Learning Objectives

After completing this exercise, the residents should:

- Be familiarized with basics of statistics and its' use;
- Identify data, variable and levels of measurements;
- Practice central tendency measures and measures of spread; and
- Represent the analyzed data in suitable graphic projections.

Part-I

Introduction

COVID-19 is a highly contagious respiratory illness caused by SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). This infectious disease first emerged in Wuhan, China, in late 2019 and quickly spread globally, leading to a pandemic. It is primarily transmitted through respiratory droplets

when an infected person coughs, sneezes, talks, or breathes. The symptoms of COVID-19 can range from mild to severe and may include fever, cough, shortness of breath, fatigue, loss of taste or smell, and body aches. As of July, 2023 there have been 760+ million cases of COVID-19 globally with nearly seven million deaths. In addition, as of late June 2023, nearly 13.5 billion doses of COVID-19 vaccines have been administered globally. Afghanistan, which was no exception during this pandemic, recorded the first case of COVID-19 in February 24, 2020 and the number of total cases has increased to more than 220,000. So far, more than 20 million doses of COVID-19 have been administered in Afghanistan [1]. Prisons have known health concerns, particularly in overcrowded and understaffed facilities. However, with occurrence of COVID-19 pandemic, prisons were considered as breeding site for the virus with immediate transmission and spread chance. The subpar living conditions and lack of access to medical facilities only overturn the situation in favor of virus spread and containment [2].

Herat province which is located in western region of Afghanistan is accounted as one of the large provinces and center of this region. Herat prison has 1450 inmates (as of July, 2022) and located city of Herat. On the 27th of July 2022 at 2:00 PM, the National Disease Surveillance and Response (NDSR) coordinator of Ministry of Public Health (MoPH) received a report from the prison's health facility in charge. The report highlighted a significant increase in Acute Respiratory Infections (ARI) among the inmates. Recognizing the urgency of the situation, the coordinator promptly shared the report with the Emergency Preparedness and Response (EPR) committee members. Realizing the gravity of the situation, the EPR committee wasted no time and quickly took action. The very next day, they assigned five teams from the Surveillance and Support Team (SST) to investigate and respond to the outbreak at the Male and Female Herat prison. The teams were equipped with the necessary resources and expertise to handle the investigation effectively. Upon their arrival at the prison, the SST teams were met with a challenging task. The outbreak had initially started on the 14th of July 2022 and had rapidly spread among the inmates. Utilizing the defined case definition and the observed signs and symptoms, such as fever, cough, shortness of breath, sore throat, diarrhea, and headaches, the teams meticulously line-listed a total of 327 cases over the course of five days. The teams promptly collected samples from all the identified cases. These samples were then sent to the renowned Herat Reference Laboratory (HRL) for further analysis and testing.

Question-1: What's Statistics and Biostatistics?

Question-2: What are the main types of statistics?

Part II

Data Collection and Line Listing

For the process of line listing, the investigation teams have been instructed to use the NDSR self-developed line list. The investigation teams' coordinator speaks up during EPR committee meeting where he emphasized on the importance of data collection process and learning about variables included in this line list.

Question-3: a) What's the definition and classification of data in terms of biostatistics?

b) What's variable and how it's classified?

Give details for both questions.

Question-4: The line list for this outbreak has been attached in the annex section of this document (Annex 1).

Review the line list, describe how many variables are there?
State the type of each variable in a table.
Identify level of measurements for each variable.

Part III

Summarizing and Analyzing Data

The assigned investigation teams arrived to the outbreak area and initiated data collection process. Upon completion of data collection line list is shared with NDSR coordinator for west region. Additionally, the public health directorate of Herat province seeks an update of the outbreak. Therefore, the coordinator must prepare a preliminary report of the situation.

Question-5: What's frequency table and its' component?

Question-6: Develop frequency table for compatible variables of the attached line list and analyze the data accordingly.

Question-7: Have an overlook to the answers of question-5 and see which variable is suitable to perform the measures of central tendency. Conduct the analysis for mean, median, mode and range on this variable and represent it in the table.

Question-8: Taking into account the age of patients, how can you classify them using the range and also develop a frequency table showing the simple and cumulative frequency with its percentages.

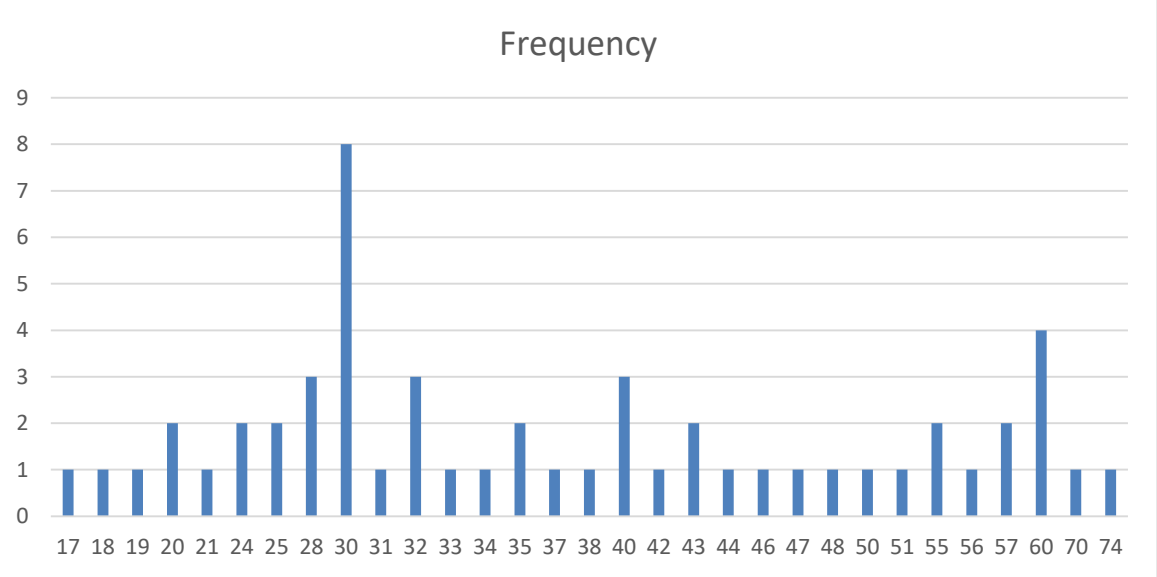
Question-9: As measure of spread, calculate the range and standard deviation of the "Age" variable manually.

Part IV

Best Use of Measures and Graphs

The basic analysis for line list has been performed however, the coordinator of NDSR is doubtful over representation of the data in his report. He has found out the mean, median, mode and standard deviation of the age data but he's not sure to use which one of them.

Question-10: For his report based on the findings, the NDSR coordinator has developed a bar chart with the following depiction. Which of the central tendency and spread measures he should use? Is data normally distributed?



Question-11: Graphically represent the suitable variables included in this line list using bar and pie chart with the help of gridline attached in annex section (Annex 2).

Part-V
Conclusion

Taking into account the importance of outbreak and letting the health authorities as well as scientific societies know about the outbreak the NDSR team has developed a brief report of one paragraph and all analysis and summarization of data were reflected.

Question-12: Write a brief report of this outbreak.

Conflict of Interest Statement

The authors declare no conflict of interest related to this manuscript.

Annex 1: Line List of COVID-19 outbreak in Herat Prison

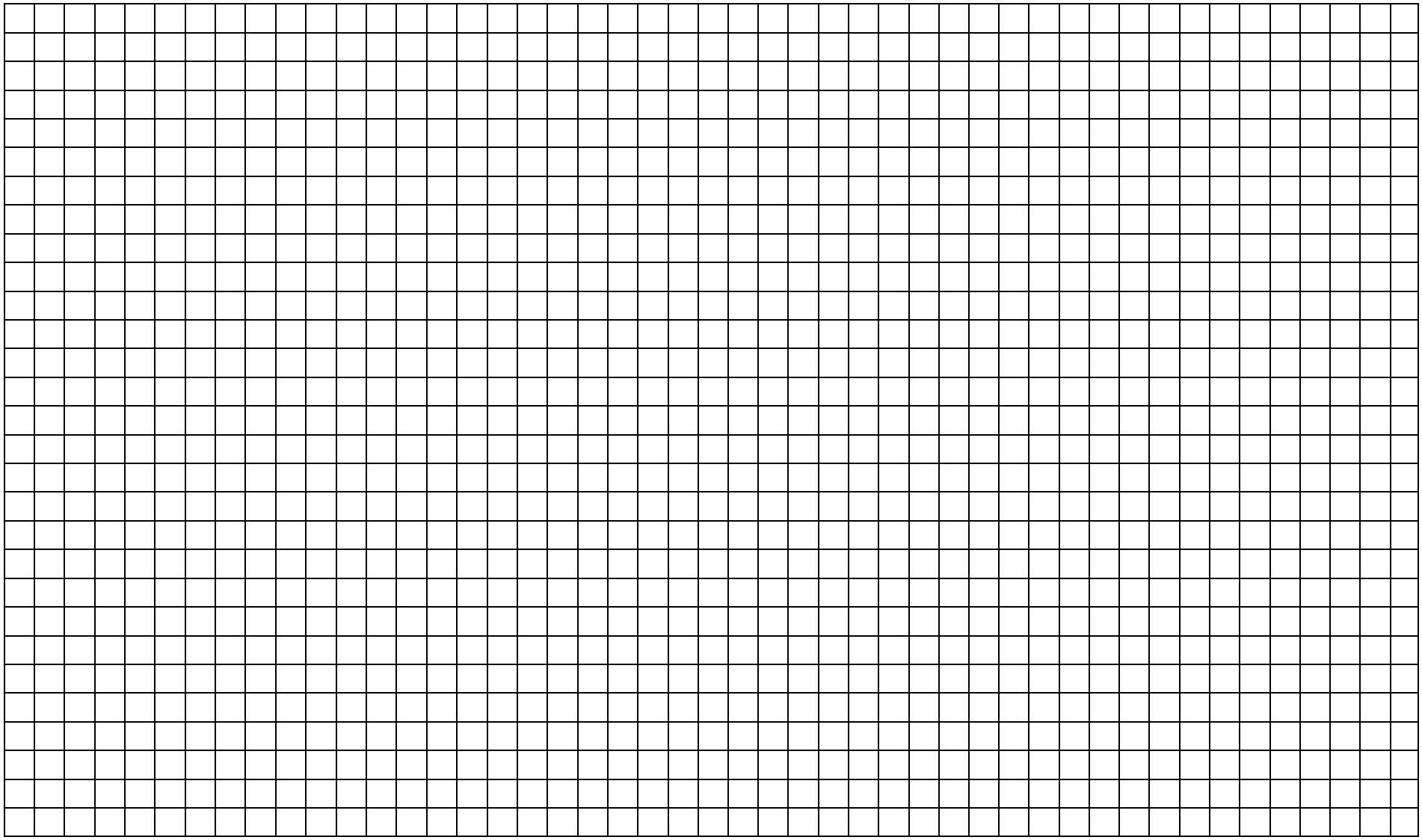
SN	Age	Gender	Sign and Symptoms	Symptoms onset	Fever	Cough	Shortness of breath	Sore throat	Diarrhea	Headache	Weakness	Test Result
1	55	Male	Yes	20-Jul-22	No	Yes	Yes	Yes	No	Yes	Yes	Positive
2	60	Male	Yes	18-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
3	43	Male	Yes	21-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
4	30	Male	Yes	21-Jul-22	Yes	Yes	Yes	No	No	No	No	Positive
5	55	Male	Yes	18-Jul-22	Yes	Yes	Yes	Yes	No	No	No	Positive
6	48	Male	Yes	15-Jul-22	Yes	No	Yes	Yes	No	Yes	No	Positive
7	60	Female	Yes	18-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
8	30	Male	Yes	20-Jul-22	Yes	No	Yes	Yes	No	Yes	No	Positive
9	35	Male	Yes	21-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
10	21	Male	Yes	20-Jul-22	No	Yes	Yes	Yes	No	Yes	No	Positive
11	33	Male	Yes	20-Jul-22	Yes	Yes	Yes	No	No	No	No	Positive
12	44	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive

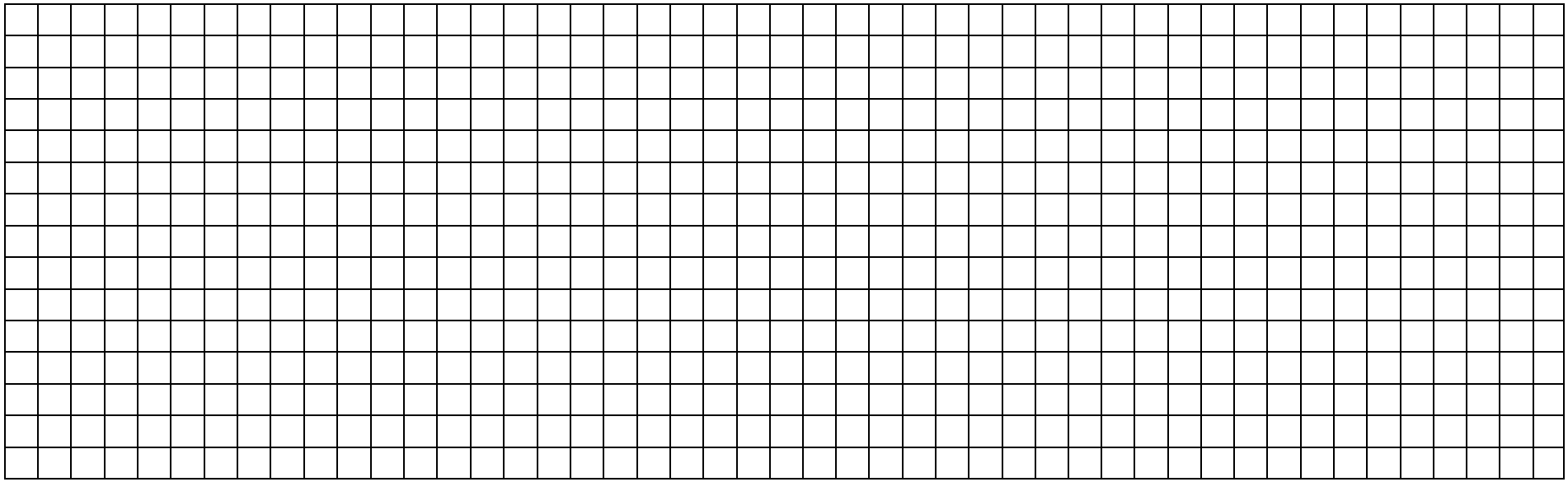
13	20	Male	Yes	15-Jul-22	Yes	Yes	Yes	Yes	Yes	Yes	No	Positive
14	25	Male	Yes	20-Jul-22	No	Yes	Yes	Yes	No	No	No	Positive
15	35	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
16	56	Female	Yes	21-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
17	24	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
18	60	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
19	70	Male	Yes	18-Jul-22	Yes	Yes	Yes	Yes	Yes	Yes	No	Positive
20	57	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
21	42	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
22	40	Female	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
23	32	Male	Yes	20-Jul-22	Yes	Yes	Yes	Yes	No	Yes	No	Positive
24	40	Male	Yes	20-Jul-22	Yes	No	Yes	Yes	No	Yes	No	Positive
25	24	Male	Yes	18-Jul-22	Yes	No	Yes	Yes	No	Yes	No	Positive
26	30	Male	Yes	20-Jul-22	Yes	Yes	Yes	No	Yes	Yes	No	Positive
27	34	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive

28	28	Female	Yes	21-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
29	30	Male	Yes	15-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
30	30	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
31	19	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
32	30	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
33	32	Male	Yes	21-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
34	60	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
35	31	Female	Yes	18-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
36	51	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
37	38	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
38	17	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
39	43	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
40	25	Female	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
41	57	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
42	18	Male	Yes	21-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive

43	32	Male	Yes	15-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
44	28	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
45	30	Female	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
46	37	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
47	30	Male	Yes	21-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
48	28	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
49	74	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
50	50	Male	Yes	18-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
51	20	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
52	47	Male	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
53	40	Male	Yes	18-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive
54	46	Female	Yes	20-Jul-22	Yes	Yes	No	Yes	No	Yes	No	Positive

Annex 2





References:

1- Who coronavirus (COVID-19) dashboard, World Health Organization. World Health Organization. Available at: https://covid19.who.int/ (Accessed: July 1, 2023).

2- Worobey M, Pekar J, Larsen BB, et al. The emergence of SARS-CoV-2 in Europe and North America. Science. 2020;370(6516):564-570. doi:10.1126/science.abc8169