Primary Objective:

- To determine the prevalence and trends of the reported FBDs cases and incidence rates of certain notifiable infectious diseases in Saudi Arabia from 2015 to 2023.

- To determine in understanding how regional factors contribute to the distribution of foodborne pathogens and aid in developing targeted interventions to improve food safety in specific areas.

- To compare foodborne outbreaks and cases caused by common sources and those caused by home sources, and evaluate the demographic characteristics.

Secondary Objectives:

1. To identify and categorize the most common Notifiable Infectious Diseases reported in Saudi Arabia during the study period.

2. To analyze the temporal distribution identifying any significant changes or trends over the years 2015-2023.

3. To appraise the frequency of microbes incident per quarter (Q1: from Jan to Mar, Q2: from April to Jun, Q3: Jul to Sep, Q4, Oct to Dec), and its association with regions

4. To map the geographic distribution of foodborne pathogen outbreaks within Saudi Arabia and identify any regional hotspots.

5. To evaluate the demographic characteristics (age, gender, nationality, etc.) of affected populations and identify vulnerable groups.

6. To investigate the seasonality of foodborne pathogen outbreaks and assess if certain pathogens are more prevalent during specific times of the year.

7. To utilize statistical methods to identify significant trends and correlations in the prevalence data of foodborne pathogens and infectious disease in all regions.

8. To compare the incidence of specific foodborne pathogens in urban versus rural areas within different regions of Saudi Arabia

Report

In this study, we have explored the prevalence and trends of notifiable diseases in Saudi Arabia from the years 2015 to 2023. Our objective was to understand how regional factors influence the types and sources of diseases, thereby providing valuable insights to improve food safety measures and public health strategies. By analyzing this patient data, we aimed to identify patterns that could help mitigate the risk of foodborne illnesses in the region.

Demographic characteristics

We demonstrate frequency and percentage for demographics variables such as gender, Age, and nationality respectively as shown in Table (1). For gender variable male have the highest value from the whole population about 8677(51.6%), where female have frequency about 8141(48.4%). In addition, for age group the highest group from 20-49 years has 7531(44.4%) responses from the total, next by age groups 5-19, 1-4 years, and +50 year have 6954 ,1440, and 957 responses, respectively. The lowest age group less 1 year has sum 74 responses. Finally, about the nationality 12681 of the respondents are Saudi, where 4275 are non-Saudi. It may be concluded that the demographic data of the patient population revealing that they belong to moderate age group and that they are male, Saudi.

Table 1. Demographics characteristics

Factor		Frequency *(n)*	Percentage *(%n)*
Gender	Male	8677	51.6
	Female	8141	48.4
	Total	16818	100
Age	less 1 year	74	0.4
	1-4 year	1440	8.5
	5-19 year	6954.	41
	20-49 year	7531	44.4
	+50 year	957	5.6
	Total	16956	100
Nationality	Saudi	12681	74.8
	Non-Saudi	42075	25.2
	Total	16956	100

Disease Prevalence and Sources

Now, we need to utilize some descriptive statistics for types diseases. The most common infection diseases are Brucellosis, Amoebic Dysentery, Salmonella, Hepatitis have total patients' cases 32091,18013,14648, and 1286 respectively. Where Bacillary Dysentery, Paratyphoid amp, and Echinococcus have total 439, 313, and 1 from whole total respectively. where each type classified into two types of source case from pubic 11610 patients, and case from hospital 5346 patients. To sum up, the most common infection diseases reported is Brucellosis, and the diseases from the public are more than diseases from hospital in Saudi Arabia reported during the study period.

Table 2. Descriptive statistics for types diseases

		Mean	Mode	Min	Max	Total
Types of diseases	Salmonella	81.83	0	740	740	14648
	Paratyphoid & Typhoid	1.75	0	0	60	313
	Amoebic Dysentery	100.63	0	0	1641	18013
	Bacillary Dysentery	2.45	0	0	37	439
	Echinococcus Hydatid	.01	0	0	1	1
	Hepatits A	7.18	0	0	70	1286
	Brucellosis	179.28	13[a]	0	782	32091
sources	case from public	64.50	0	0	349	11610
	case from hospital	29.70	0	0	230	5346

Trends Over Time

To analyze the trends of the mean distribution for infectious diseases over the years from 2015-2023 we use time series as shown in Figure1. We notice that from 2018 into 2019, the mean of Salmonella, Bacillary, Hepatitis A, and Brucellosis rapidly increasing, but these types of infectious diseases decreasing in two years 2020 and 2022. Where, the mean of paratyphoid & Typhoid and Echinococcus have the same trend drop down rapidly in 2017, and 2016 to go into zero mean respectively. Finally, the mean of Amoebic Dysentery slowly increasing from 2016 into 2017, from 2017 into 2020 stably decreasing, and from 2020 into 2021 constant. Now, we discover the mean distribution of different infectious diseases over Q1, Q2, Q3, and Q4 as shown in Figure2. From Q2 into Q3 there are regular increasing in the mean of Salmonella,Paratyphoid &amp, Typhoid, Amoebic Dysentery, Bacillary Dysentery, and Brucellosis. In addition, the mean of these types of infectious diseases Paratyphoid & Typhoid and Bacillary Dysentery from Q3 into Q4 still increasing. But, Salmonella, Amoebic Dysentery, and Brucellosis decreasing. Finally, Echinococcus Hydatid have triangle shape with rapidly increasing from Q1 into Q2, then rapidly decreasing from Q2 into Q3.

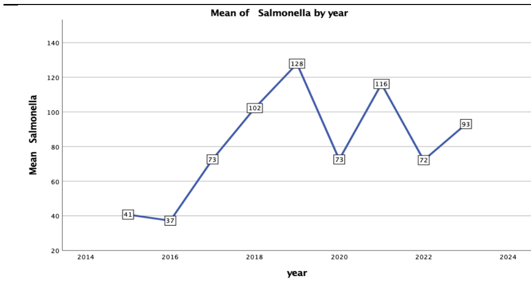


Figure 1.1

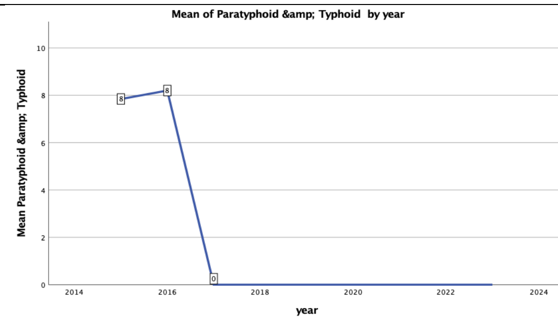


Figure 1.2

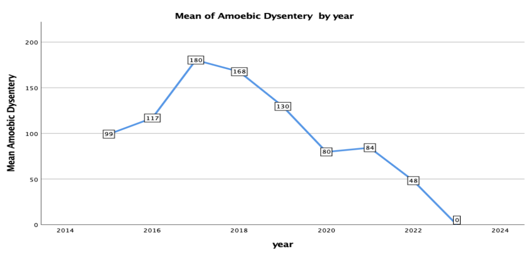


Figure 1.3

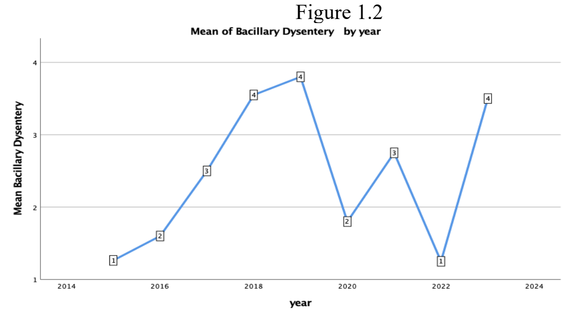


Figure 1.4

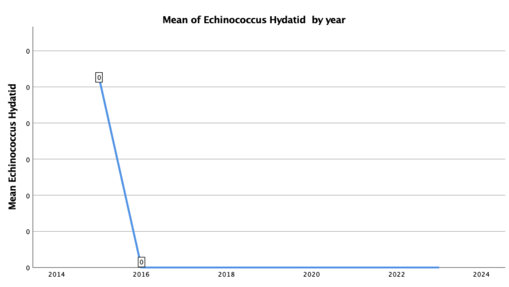


Figure 1.5

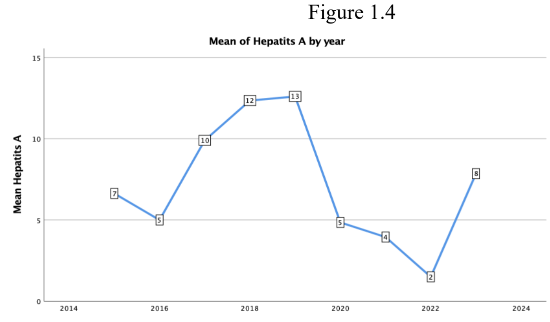


Figure 1.6

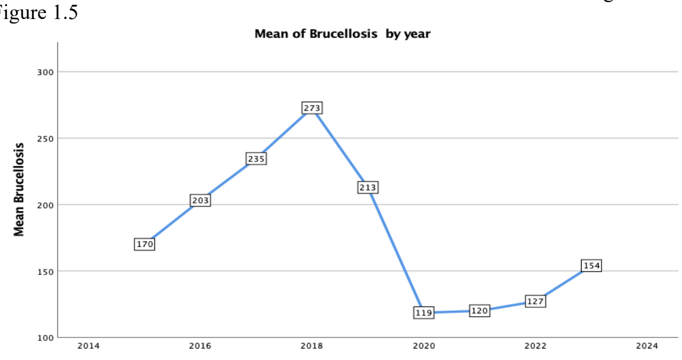


Figure 1.7

Figure1. Time series for the mean distribution of infectious disease over year in Saudi Arabia.

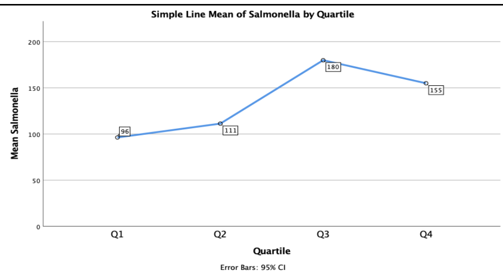


Figure 2.1

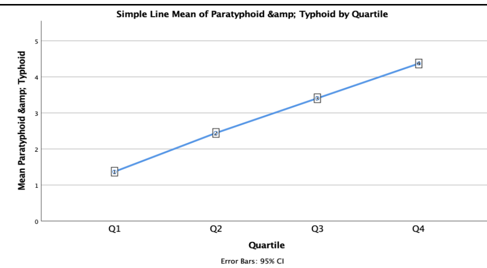


Figure 2.2

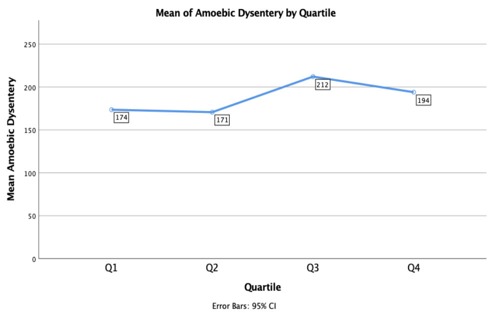


Figure 2.3

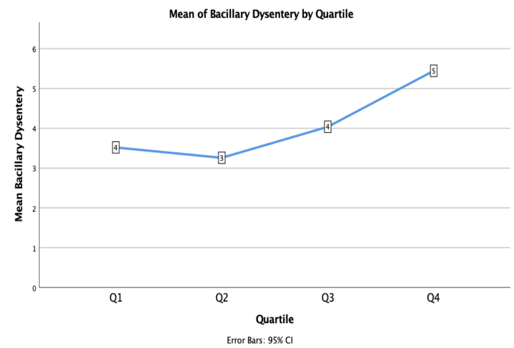


Figure 2.4

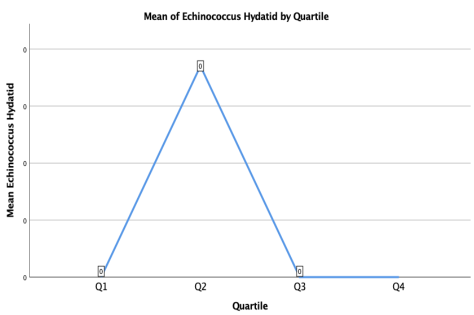


Figure 2.5

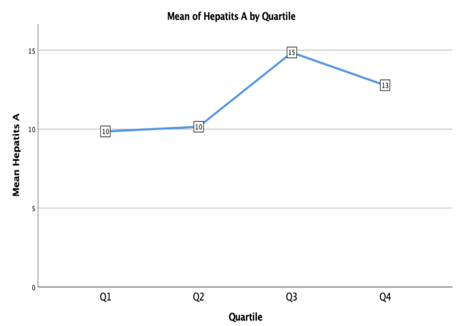


Figure 2.6

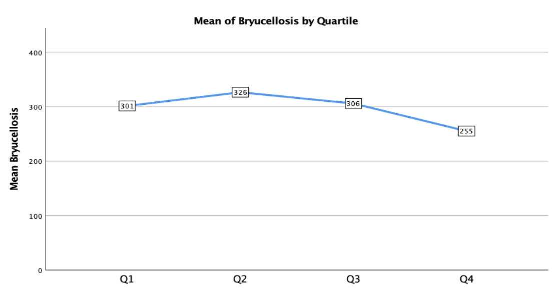


Figure 2.7

Figure2. Time series for the distribution of infectious disease over Quartile in Saudi Arbia

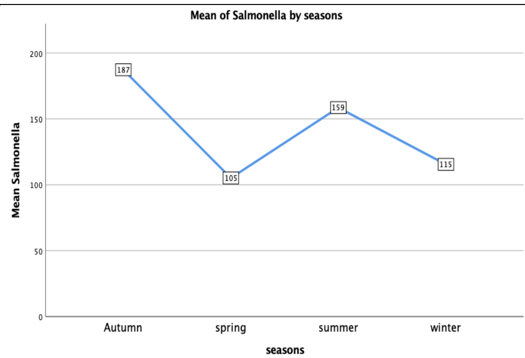


Figure 3.1

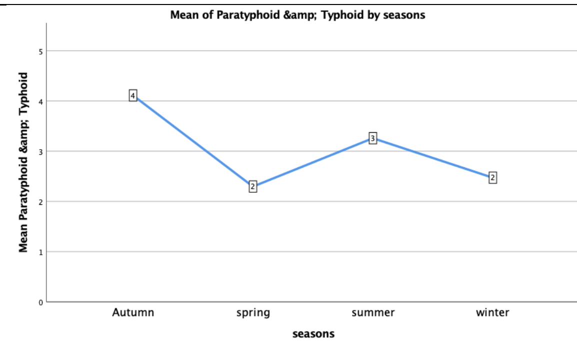


Figure 3.2

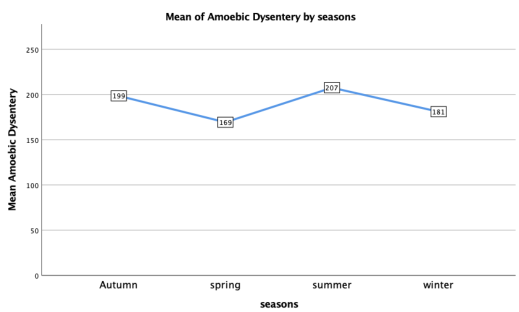


Figure 3.3

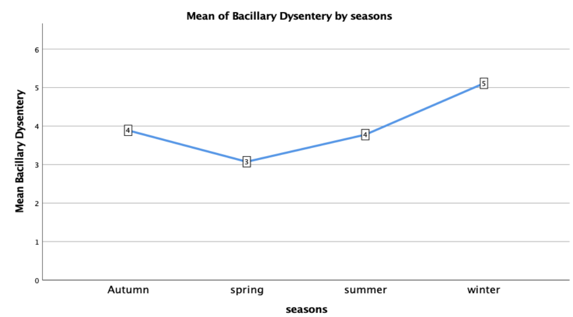


Figure 3.4

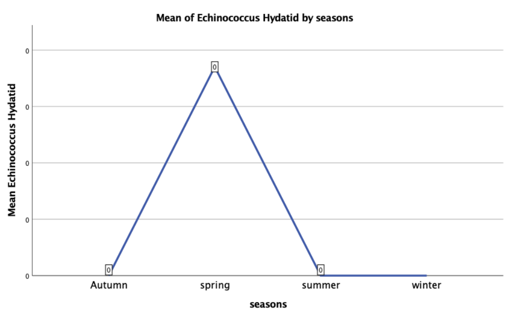


Figure 3.5

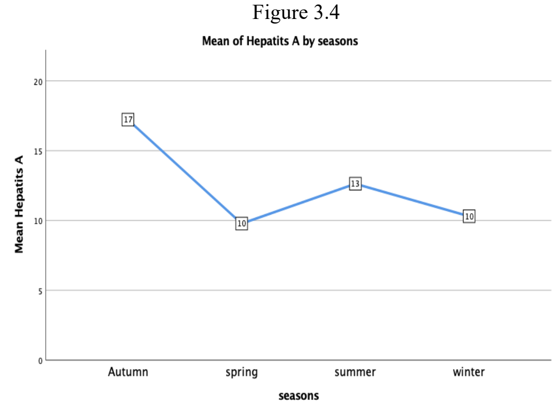


Figure 3.6

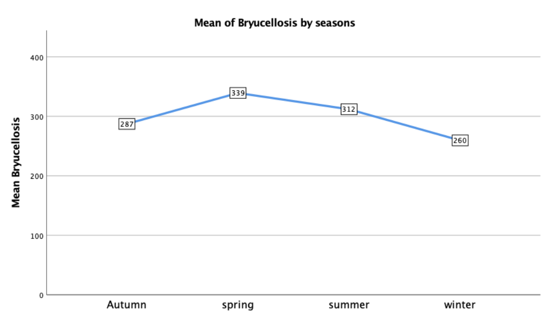


Figure 3.7

Figure3. seasonality time series distribution of infectious disease in Saudi Arbia

To investigate the mean distribution of the different infectious diseases over the seasons in the year. In general, we notice that Salmonella, Amoebic Dysentery, Paratyphoid & Typhoid, and Hepatits A are decreasing from Autumn into spring and from summer into winter, but decreasing from spring to summer. In addition, Bacillary Dysentery has the same trend from Autumn into summer, bycontrast there is increasing trend from summer into winter. In the other side, Brucellosis has increasing trend from Autumn into spring, then decreasing from spring into winter . Finally, Echinococcus Hydatid has different trend shape like triangle rapidly increasing from Autumn into spring, then rapidly decreasing from spring into summer with zero mean from summer into winter.

To sum up, there are a lot of types of infectious diseases grownup from spring into summer as shown in Figure3.

From Table3 and Figure 3 , we can identify the association between regions and the types of diseases where Region is coded from 1-19 for Riyadh to Quaryyat consequency. Where each increasing in coding for region decreasing in the mean of infectious diseases. For Hepatitis A and Brucellosis about .585 and .505 respectively, its mean negative strong correlation. In addition, each increasing in coding for region decreasing Salmonella, Amoebic Dysentery, and Bacillart Dysentery about .478,.479, and .375 respectively, its means weak negative correlation. Finally, there is no correlation between the region and Paratyphoid & amptyphoid , Echinococcus consequently where sig>0.00.

Table 3. correlation between types diseases and region

Spearman's rho		
Salmonella*Region.	Sig	.000
	Correlation	-.478**
Paratyphoid &amptyphoid*Region.	Sig	.106
	correlation	-.125
Amoebic Dysentery *Region	Sig	.000
	Correlation	-.375**
Bacillary Dysentery	sig	.000
	correlation	-.479**
Echinococcus	sig	.576
	Correlation	-.044
Hepatitis A	sig	.000
	correlation	-.585**
Brucellosis	sig	.000
	Correlation	-.505**

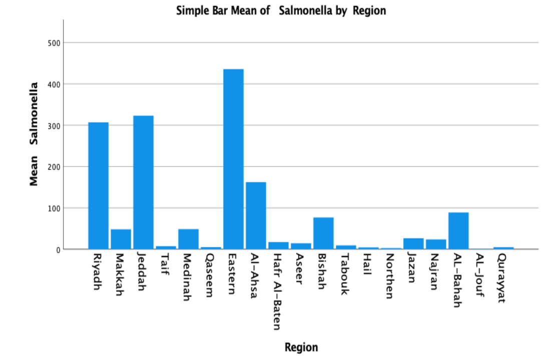


Figure4.1

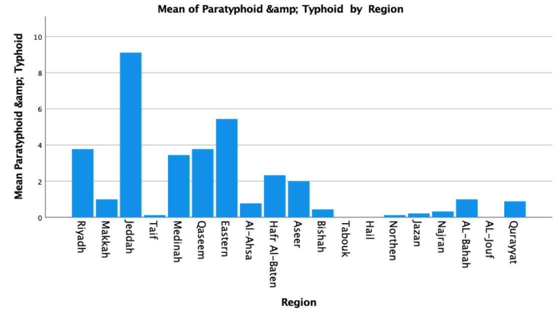


Figure 4.2

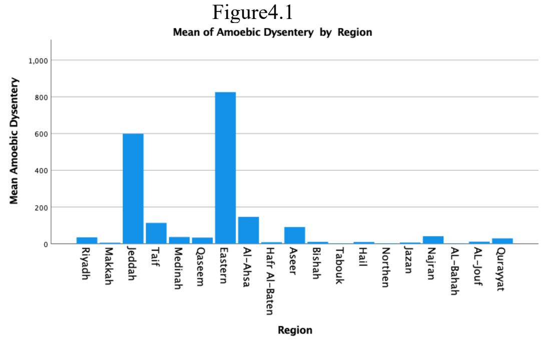


Figure 4.3

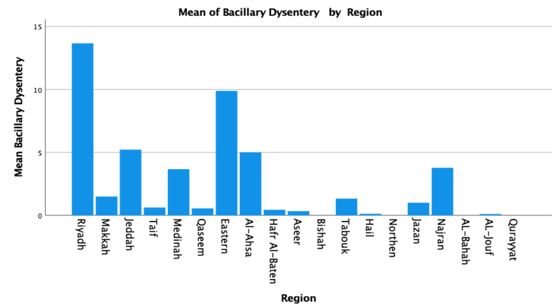


Figure 4.4

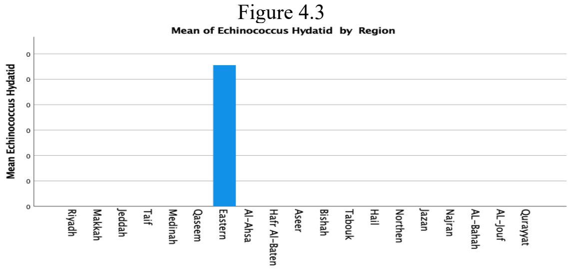


Figure 4.4

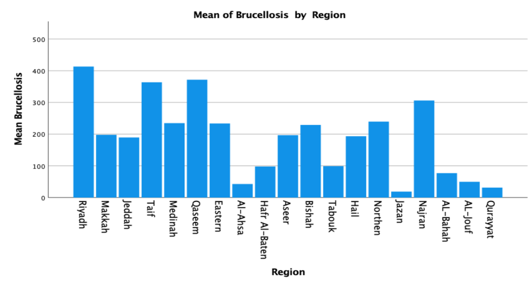


Figure 4.5

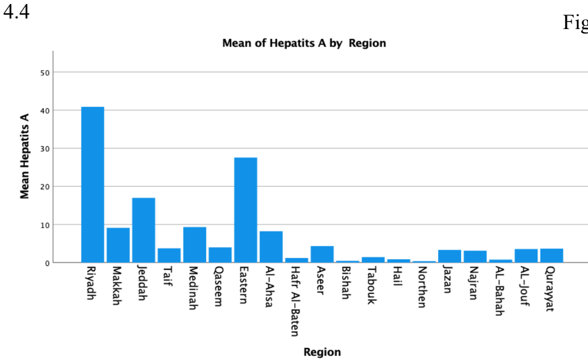


Figure 4.6

Figue3. Bar chart distribution of infectious disease in different region in Saudi Arbia

Urban vs. Rural Comparison

By using Kruskal -Wallis H to identify significant different in the mean of each infectious disease between all the region. We conclude there are difference in the mean of Salmonella, Amoebic Dysentery, Bacillary Dysenttery, Echinococcus , Hepatitis A, and Brucellosis(sig<.00) in each region in Saudi Arabia ,where is no difference between the regions in the mean of Paratyphoid (sig>.00).

Table 3. correlation between types diseases and region using Kruskal Wallis Test

	Paratyphoid	Salmonella	Amoebic Dysentery	Bacillary Dysentery	Echinococcu	Hepatits A	Brucellosis
Kruskal-Wallis H	9.189	102.904	66.229	110.222	17.556	99.498	106.821
df	18	18	18	18	18	18	18
Asymp. Sig.	.955	<.000	<.000	<.000	.485	<.000	<.000

Table4. Mantiney test for difference in the mean of disease between rural and urban .

	Salmonella	Paratyphoid & Typhoid	Amoebic Dysentery	Bacillary Dysentery	Echinococcus Hydatid	Hepatits A
Mann-Whitney U	495.500	703.000	608.000	548.500	760.500	706.500
Wilcoxon W	540.500	15238.000	15143.000	593.500	805.500	15241.500
Z	-1.784	-.638	-1.053	-1.557	-.230	-.393
Asymp. Sig. (2-tailed)	.074	.523	.292	.119	.818	.695

By using Mann-Whitney test, we notice that there is no difference in the mean of all diseases between rural and urban region in Saudi Arbia.

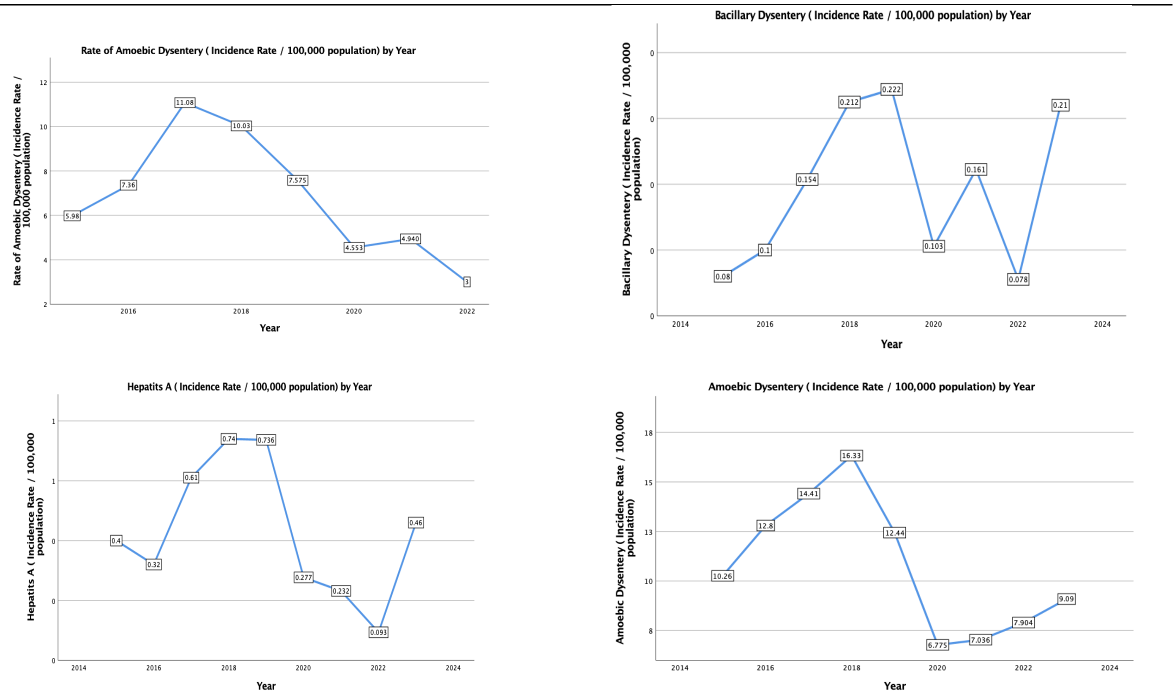


Figure4. Time series changes the infectious rate over the years

Figure4 shows time series changes of infectious rate over all years. Where the infectious rate generally increasing from 2017-2018.In addition prevalence rate for Saudi Arabia for Bacillary Dysentery and Hepatitis are increasing from 2018 into 2019, for Amoebic Dysentery is increasing from 2017 into 2018, for Brucellosis in 2018 as shown Figure 5.

Figure5 shows that prevalence rate for non-Saudi Arabia for Amoebic Dysentery, bacillary Dysentery, Hepatitis, and Brucellosis respectively. Where Amoebic Dysentery has highest prevalence rate in 2017, for bacillary Dysentery prevalence rate increasing regularly, for Hepatitis prevalence rate nearly constant, and for Brucellosis prevalence rate decreasing carve.

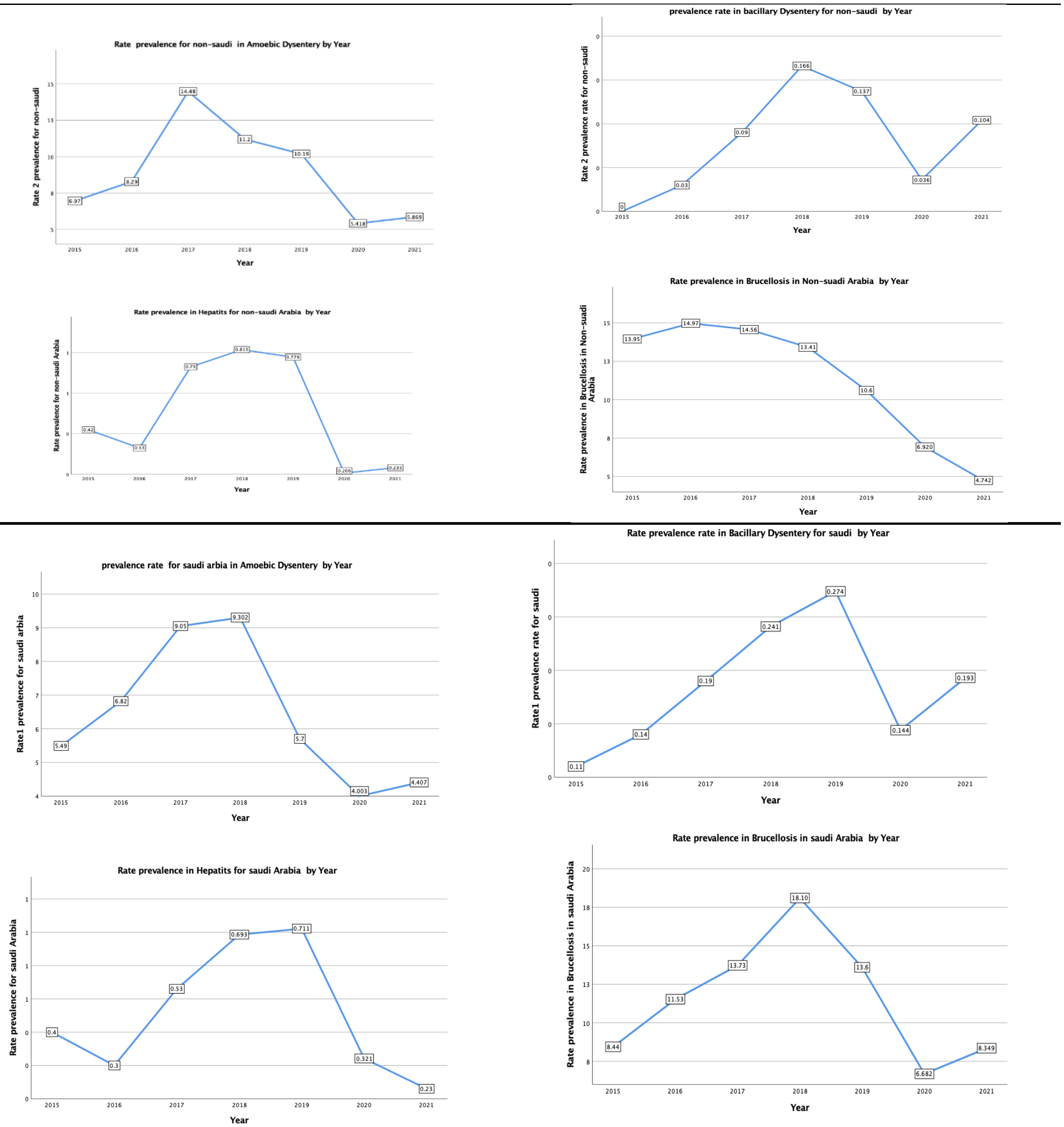


Figure5. Time series changes the prevalence rate for Saudi Arabia and non-Saudi Arabia over the years

Table 5 displays the results of the linear regression analysis that was done to determine whether the age in each group less than one, from 1-4 , from 5-19, from 20-49,and more than 50 years on the dependent variable cases from public source. As can be seen from the table, the regression coefficient (B) is equal to -2.085, each age group from one to more than 50 significantly effect on the number of cases from public source, ($p = 0.005$, significant at $p < 0.01$). For age group (1-4) years each increase for one unit will decrease the number of cases about one unit ,for age group(5-19) years each increase for one unit will increase the number of cases about .8, for age group(20-49)years each increase for one unit will increase the number of cases in public source about one unit , and for age group more than 50 years each increase for one unit will increase the number of cases in public about .620. In general, the age group effect on the variety of the number of cases in public source about 95%

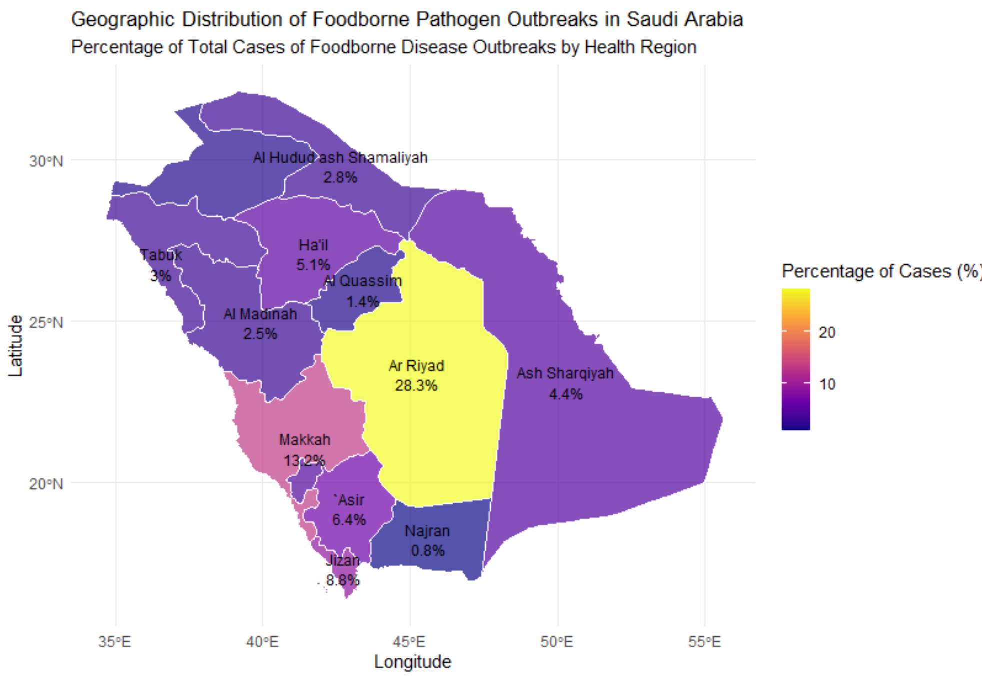
Table5.factor effect on the cases from public source

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-2.085	2.488		-.838	.403
less 1-year	-1.574	2.280	-.017	-.690	.491
from 1-4 year	-1.368	.323	-.143	-4.238	.000
from 5-19 year	.818	.071	.417	11.455	.000
from 20-49	1.035	.043	.687	23.870	.000
more 50 year	.620	.120	.122	5.155	.000

Table6. R square for regression equation

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.958[a]	.917	.915	22.521

To map the geographic distribution of foodborne pathogen outbreaks within Saudi Arabia and identify any regional hotspots.



Results and Discussion

Conducting the analysis of Saudi Arabia foodborne pathogen outbreak from 2015 to 2023 has provided valuable information concerning the prevalence, geographic distribution, and diseases patterns. The analysis has taken into consideration various aspects including assessing the lakoues of nutrition taboot of the subjects involved, their geographic distribution, estimated populations and the age distribution of the infected persons. According to the descriptive statistics, the most common infectious diseases in the outbreak included Brucellosis, Amoebic dysentery, Salmonella, and Hepatitis A in an increasing order of their prevalence. The lakoue of nutrition taboot totalled to 35789.000 in Brucellosis which was the most common foodborne pathogen with 32091 camoues, Amoebic dysentery had 18013 cases, Shigella 14648 cases and Hepatitis A had 1286 cases. The distribution of the cases was as follows, more cases were referred from the public sector 11610 cases in comparison to the hospitals which had 5346 cases indicating some difficulties in the public health in the settings.

The time series analysis illustrates different trends in the prevalence of these diseases across the years. For example, based on the annual mean distribution, Salmonella, Bacillary Dysentery, Hepatitis A, and Brucellosis increased rapidly from 2018 to 2019 and declined in 2020 and 2022. However, Paratyphoid & Typhoid and Echinococcus Hydatid decreased quickly from 2017 and 2016, respectively, to a zero mean. Conversely, the seasonal analysis demonstrates some differences. Salmonella and Amoebic Dysentery grew from Q2 to Q3, whereas, Echinococcus Hydatid is a triangular wave with a peak of Q2 and quickly decreases to Q3.

Regarding the correlation analysis of disease types and regions, the results showed a strong negative correlation of Hepatitis A and Brucellosis of the regional coding, meaning that the higher the regional code, the lower is the prevalence of the diseases. Also, weak negative correlations were found for Salmonella, Amoebic Dysentery, and Bacillary Dysentery. There was no significant correlation found between regions and Paratyphoid & Typhoid and Echinococcus.

Saudi Arabia requires targeted public health interventions to combat foodborne pathogen outbreaks by the emphasis on the regional and seasonal characteristics. The high prevalence of such diseases as Brucellosis in public settings points to the necessity of community-based public health forums and education programs as being essential to curb the rates of disease transmission. The significant drop in the rates of the target diseases during actual years is supposed to show that the relevant public health interventions were somewhat effective. However, the reemergence of the given diseases in following years shows that further measures must be adopted to control the target diseases more effectively.

Overall, the reported case investigation and metagenomic analysis of a foodborne outbreak in 2014 in Saudi Arabia. The results of the study indicate the importance of more effective public health policies and strategies being developed in order to reduce the burden of these diseases in a region. I think that region and season-specific features underscore the need for a closer look at underlying factors. The current level of intervention in the report did not offer any conclusive results but I believe that new intervention methods should be considered.

Conclusion

This comprehensive analysis presents meaningful information on the distribution and lines of development of infectious diseases taking place in Saudi Arabia within recent 8 years. To a large extent, it shows that, in such a region, regional and seasonal specificities are crucial in terms of the incidence. Also, the present study proves that in such circumstances, the government is to implement the particular intervention.

References

1. Alshammari, M., Alzahrani, H., & Al-Asmari, M. (2019). Epidemiology of foodborne diseases in the Eastern Province of Saudi Arabia. Journal of Infection and Public Health, 12(6), 861-867. https://doi.org/10.1016/j.jiph.2019.05.005
2. Al-Tawfiq, J. A., & Memish, Z. A. (2019). Emerging respiratory and foodborne infectious diseases: A systematic review. International Journal of Infectious Diseases, 87, 44-51. https://doi.org/10.1016/j.ijid.2019.07.010
3. Bin Saeed, A. A., Al-Hamdan, N. A., & Fontaine, R. E. (2021). Epidemiological trends of Hepatitis A in Saudi Arabia: A longitudinal study. Journal of Public Health, 43(3), 506-513. https://doi.org/10.1093/pubmed/fdaa145.
4. El-Gamal, M. I., Anbar, H. S., & El-Azab, A. S. (2018). Foodborne pathogens in Saudi Arabia: Incidence, antimicrobial resistance, and prevention. Journal of Infection and Public Health, 11(6), 727-732. https://doi.org/10.1016/j.jiph.2018.07.007.
5. Hussein, R. A., & Mohamed, S. A. (2017). The impact of public health interventions on the incidence of Bacillary Dysentery in Saudi Arabia. Saudi Journal of Medicine & Medical Sciences, 5(4), 312-318. https://doi.org/10.4103/sjmms.sjmms_34_17.
6. Khan, M. A., Al-Humaidi, R. B., & Al-Ghanem, H. M. (2020). Seasonal variation and regional distribution of foodborne illnesses in Saudi Arabia. Eastern Mediterranean Health Journal, 26(8), 974-981. https://doi.org/10.26719/emhj.20.049.

Appendix

```
# R-code  for Map

##

######

# Install necessary libraries if not already installed

# install.packages(c("ggplot2", "dplyr", "sf", "rnaturalearth", "rnaturalearthdata", "viridis"))



# Load necessary libraries

library(ggplot2)

library(dplyr)

library(sf)

library(rnaturalearth)

library(rnaturalearthdata)

library(viridis)



# Read the data from the CSV file

data <- read.csv("C:/Maryam_Project_30th_July_2024/fwddataanalysisresultsection/data1.csv")
```

```
##########################

######

# Set working directory

setwd("C:/Maryam_Project_30th_July_2024/fwddataanalysisresultsection/")



# Install necessary libraries if not already installed

# install.packages(c("ggplot2", "dplyr", "sf", "rnaturalearth", "rnaturalearthdata", "viridis"))

# Load necessary libraries

library(ggplot2)

library(dplyr)

library(sf)

library(rnaturalearth)

library(rnaturalearthdata)

library(viridis)



# Read the data from the CSV file

data <- read.csv("data1.csv")



# Calculate the total number of cases
```

```r
total_cases <- sum(data$Total.Cases.of.FBDOs)


# Calculate the percentage of total cases for each health region

data <- data %>%

  mutate(Percentage.Cases = (Total.Cases.of.FBDOs / total_cases) * 100)


# Download the shapefile for Saudi Arabia's administrative regions

saudi_shapefile <- ne_states(country = "Saudi Arabia", returnclass = "sf")


# Inspect the column names of the shapefile to find the correct column for regions

names(saudi_shapefile)


# Merge the data with the shapefile

# Assuming 'name' is the column in the shapefile and 'Health.Region' in your data

saudi_data <- saudi_shapefile %>%

  left_join(data, by = c("name" = "Health.Region"))


# Plot the map

ggplot(saudi_data) +
```

```
  geom_sf(aes(fill = Percentage.Cases), color = "white", alpha = 0.7) +

  scale_fill_viridis_c(option = "plasma", name = "Percentage of Cases (%)") +

  theme_minimal() +

  labs(title = "Geographic Distribution of Foodborne Pathogen Outbreaks in Saudi Arabia",

       subtitle = "Percentage of Total Cases of Foodborne Disease Outbreaks by Health Region",

       x = "Longitude",

       y = "Latitude") +

  theme(legend.position = "right") +

  geom_sf_text(aes(label = paste0(name, "\n", round(Percentage.Cases, 1), "%")), size = 3, check_overlap = TRUE)

##########################################
```