Predictive Healthcare Analysis of Elevens Disease Categories

Waree Protprommart
CSCI E-116

Project purpose and motivation

Objective

When we look back at human history, there was a period when some diseases were on the rise. Imagine how beneficial it would be to the human race if we could detect the trends or seasonality of the disease. In this project, I aim to conduct a United States data time series analysis of eleven disease categories, including Autoimmune, Bacterial, Cardiovascular, Chronic, Genetic, Infectious, Metabolic, Neurological, Parasitic, Respiratory, and Viral, starting from the year 2000 until 2024 by using the ARIMA model. Afterward, I also aim to predict the mortality rate in all disease categories by using the other 18 features using a dynamic panel data model. The features for predicting disease mortality include Country, Prevalence rate, Incidence rate, Age group, Gender, Population Affected, Health care Access, Doctors per 1000, Hospital Beds per 1000, Treatment type, Average treatment cost, Availability of vaccine treatment, Recovery rate, DALYs (Disability-Adjusted Life Years, a measure of disease burden), Improvement in 5 years, per capita income, Education index, Urbanization rate.

Significance

The significance of being able to detect the disease lies in the fact that medical staff can be prepared to allocate resources to the particular category of disease. By having a predictive healthcare system, we can lower the mortality rates of specific diseases through preventative measures. In the United States, the cost of healthcare is expensive, but the life expectancy is less than in other countries that have more affordable healthcare (Nolte E, 2008). A healthcare system focused more on diagnosis than prevention would not do well at the large population level and would not encourage the general public to maintain a healthy and good quality of life. The concept of predictive health consists of defining health as a whole experience that is not mere absence of disease, unhealth is defined as any deviation from health, and applying quality and cost-effective intervention to improve quality of life in a whole population (Brigham, 2010).

Information of the Dataset

This dataset is synthetic data that is based on the trends of real-world global health big data. It was created by Malaiarasu on Kaggle. There are a total of 22 columns in this dataset with 1000000 observations. The description of the columns is as follows:

Column name	**Description**
Country	The name of the country where the health data was recorded.
Year	The year in which the data was collected.
Disease Name	The name of the disease or health condition tracked.
Disease Category	The category of the disease (e.g., Infectious, Non-Communicable).
Prevalence Rate (%)	The percentage of the population affected by the disease.
Incidence Rate (%)	The percentage of new or newly diagnosed cases.
Mortality Rate (%)	The percentage of the affected population that dies from the disease.
Age Group	The age range most affected by the disease.
Gender	The gender(s) affected by the disease (Male, Female, Both).
Population Affected	The total number of individuals affected by the disease.
Healthcare Access (%)	The percentage of the population with access to healthcare.
Doctors per 1000	The number of doctors per 1000 people.
Hospital Beds per 1000	The number of hospital beds available per 1000 people
Treatment Type	The primary treatment method for the disease (e.g., Medication, Surgery).
Average Treatment Cost (USD)	The average cost of treating the disease in

	USD.
Availability of Vaccines/Treatment	Whether vaccines or treatments are available.
Recovery Rate (%)	The percentage of people who recover from the disease.
DALYs	Disability-Adjusted Life Years, a measure of disease burden.
Improvement in 5 Years (%)	The improvement in disease outcomes over the last five years.
Per Capita Income (USD)	The average income per person in the country.
Education Index	The average level of education in the country.
Urbanization Rate (%)	The percentage of the population living in urban areas.

The year in which the data was collected is in the range of 2000 to 2024. There are 11 categories of disease as previously mentioned, and 20 countries including: Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Mexico, Nigeria, Russia, Saudi Arabia, South Africa, South Korea, Turkey, the UK, and the USA.

Visualizations

Exploratory Data Analysis

First, let's take an initial look at the trends in each disease category in the United States in Figure 1. We can see that most of the disease categories are fluctuating around the mean. It is not clear if there is a clear trend, but there is an upward trend in the duration of 24 years for the mean mortality rate from viral disease in the United States.

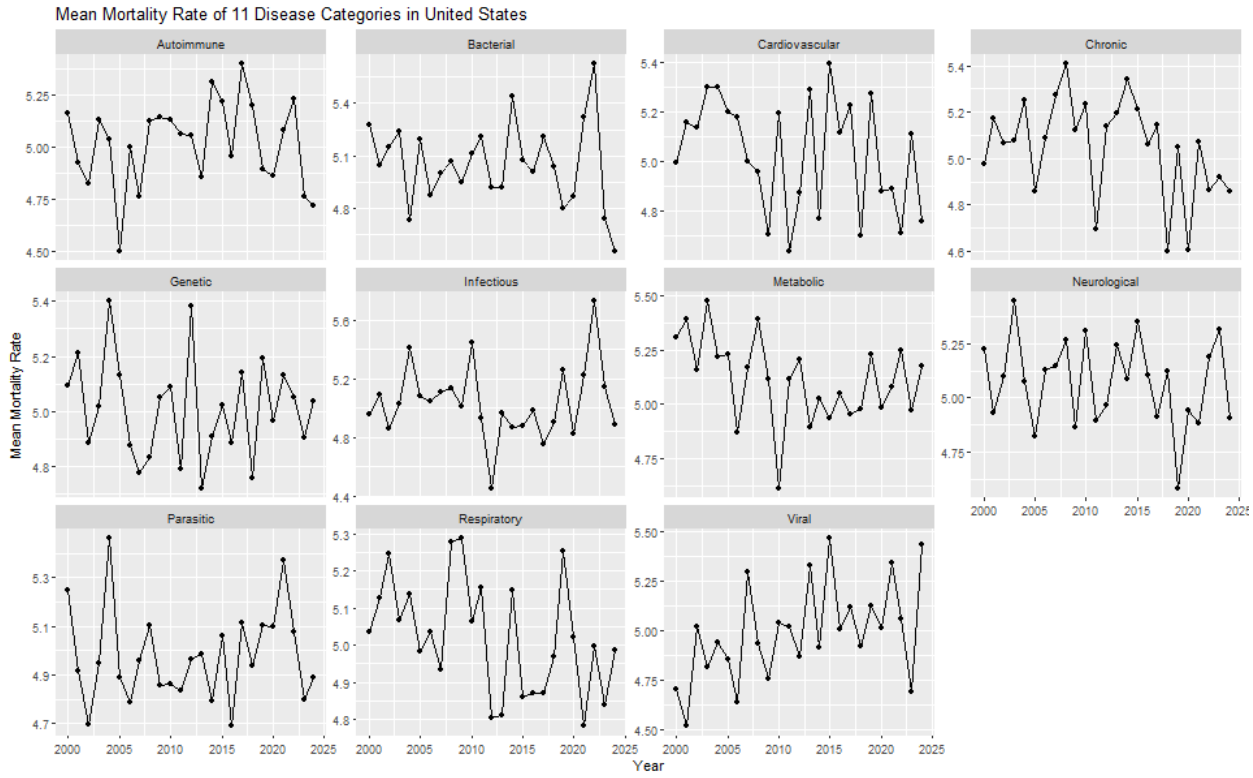


Figure 1. Mean mortality rate of eleven disease categories from 2000 until 2024.

When looking at the mortality rates of each disease category amongst the 20 different countries, each country has different trends and fluctuations of the mortality rate within the same disease category (Figure 2 shows the mean mortality rate by cardiovascular and viral disease as an example). For example, India has a sudden drop in the cardiovascular category around the year of 2013 while other years are relatively high. From Figure 1, we discussed that the United States has an upward trend in the mean mortality rate from viral disease. If we compare the United States time series with other countries, we see some other countries that fluctuate around the mean, like Australia, Germany, Turkey, etc. This shows that the mortality rate of a particular disease trend observed in the United States does not translate to the same observation for the rest of the world.

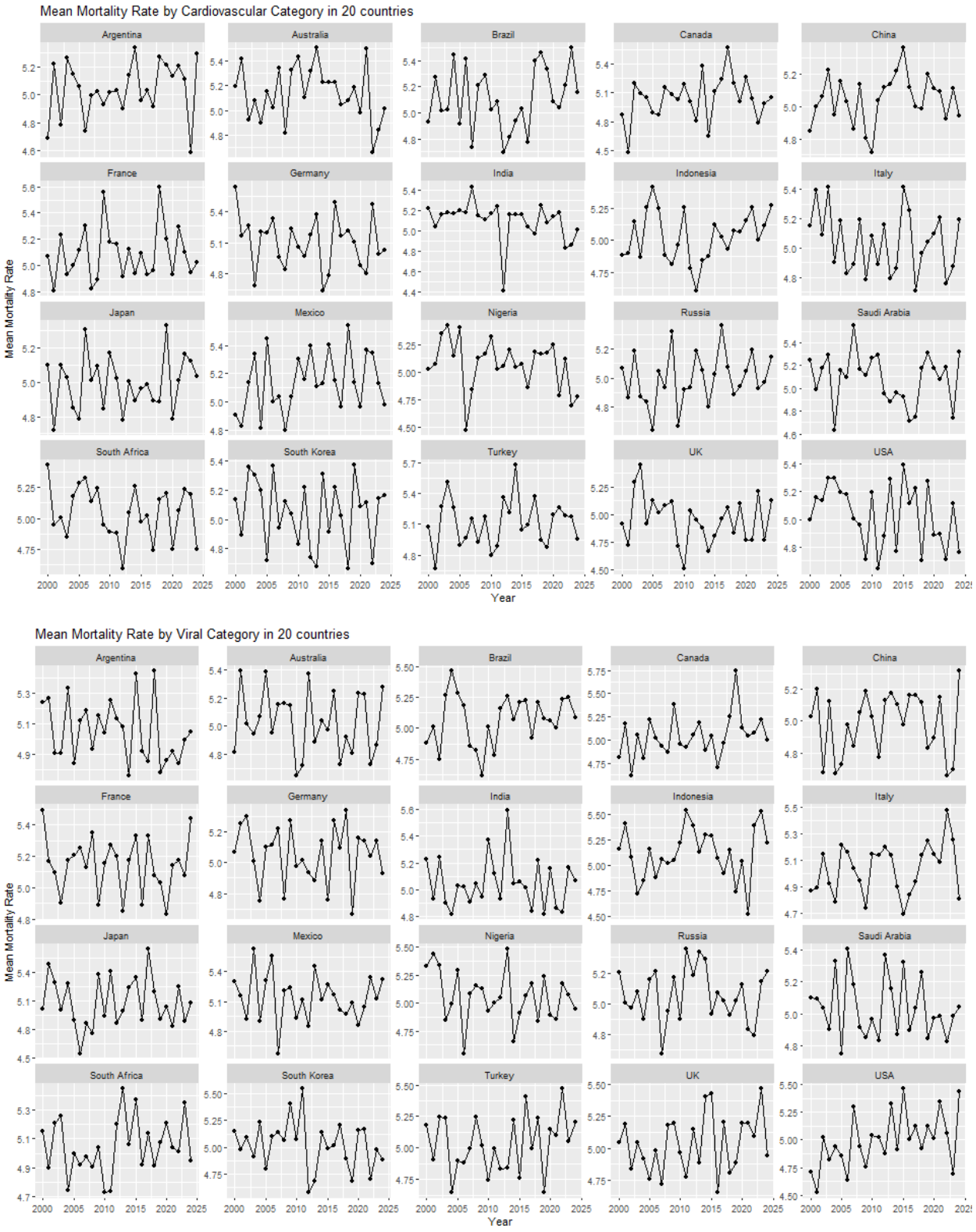


Figure 2. Mean mortality rate by cardiovascular and viral disease amongst twenty countries from 2000 until 2024. More disease categories are available in the appendix.

ARIMA model

Model Selection and Justification

The first objective of this project is to analyze any trends or seasonality from its own time series data of eleven disease categories for the United States. Autoregressive Integrated Moving Average or ARIMA is suitable for this purpose because it composed of the three components including **autoregression (AR)**, which models the influence of past values on the current value; the **Integrated (I)** component, which accounts for differencing needed to make the series stationary; and the **moving average (MA)** component, which captures the influence of past forecast errors or shocks. Furthermore, ARIMA allows for the investigation of only by using the timeseries of mean mortality rate by itself without the external predictors involved, as we are aiming to investigate if it's possible to predict the future of the disease's mortality rate by its past data. Before beginning to fit the time series to the ARIMA model, the data needs to be converted to a stationary time series. The table below shows the number of differencing needed for the series to be stationary, along with the Augmented Dickey-Fuller Test p-value. The significance of a p-value lower than 10% or 5% shows that the series rejects the null hypothesis of the presence of a unit root and a non-stationary series. All of the time series are under 10% for the ADF p-value.

Disease Category	**Difference lag(s)**	**Augmented Dickey-Fuller Test p-value**
Autoimmune	1	0.02177
Bacterial	1	0.01
Cardiovascular	1	0.04351
Chronic	4	0.05476
Genetic	1	0.02841
Infectious	2	0.03271
Metabolic	1	0.01
Neurological	1	0.04762
Parasitic	1	0.01
Respiratory	4	0.01666

Viral	1	0.01

After fitting all the time series to the ARIMA model through the function auto. In R, the result of the p (order of regression), d (order of differencing), and q (order of moving average) components that are most suitable for each time series is stored in the table below. Disease categories of autoimmune, bacterial, metabolic, parasitic, and respiratory resulting models have the structure of (0,0,0) for p,d, and q components. This indicates that these time series are behaving like white noise after differencing, and there are no autoregressive and moving average terms present. This time series cannot use its own past value or shocks to determine current or future values. On the other hand, there are some disease categories that contain autoregressive or moving average terms. Chronic, Genetic disease, and Infectious contain the order regression of 1, 1, and 2, respectively. Cardiovascular, Neurological, and Viral contain the size of moving average windows of 2, 1, and 2. From this step, the ARIMA model has given us information about which disease category is hard to predict from only the past data, and which one is able to be structurally modeled based on past values or shocks.

Disease Category	**ARIMA model (p,d,q)**
Autoimmune	ARIMA(0,0,0) with zero mean
Bacterial	ARIMA(0,0,0) with zero mean
Cardiovascular	ARIMA(0,0,2) with zero mean
Chronic	ARIMA(1,0,0) with zero mean
Genetic	ARIMA(1,0,0) with zero mean
Infectious	ARIMA(2,0,0) with zero mean
Metabolic	ARIMA(0,0,0) with zero mean
Neurological	ARIMA(0,0,1) with zero mean
Parasitic	ARIMA(0,0,0) with zero mean
Respiratory	ARIMA(0,0,0) with zero mean
Viral	ARIMA(0,0,2) with zero mean

Out-of-sample Validation

The next step after fitting the ARIMA model to the time series is to forecast the mean mortality rate in each disease category and calculate the RMSE from the real test time series. The train dataset is from 2001 to 2021, and the test dataset is from the year 2022 to the year 2024. The table below lists RMSE values of each disease category in order from lowest to highest RMSE value. Among the disease categories that contain autoregressive or moving averages terms, Cardiovascular disease is doing the best at predicting the mean mortality rate, while chronic and infectious diseases are doing the worst at predicting the mean mortality rate. Viral and Genetic is performing better than the Neurological disease category at predicting the mean mortality rate. Other disease categories that do not contain autoregressive or moving averages always predict the mean mortality rate as a constant mean value or zero after differencing.

Disease Category	RMSE values
Cardiovascular	0.12
Bacterial	0.15
Respiratory	0.20
Autoimmune	0.23
Metabolic	0.23
Viral	0.27
Genetic	0.27
Parasitic	0.27
Neurological	0.33
Chronic	0.56
Infectious	0.56

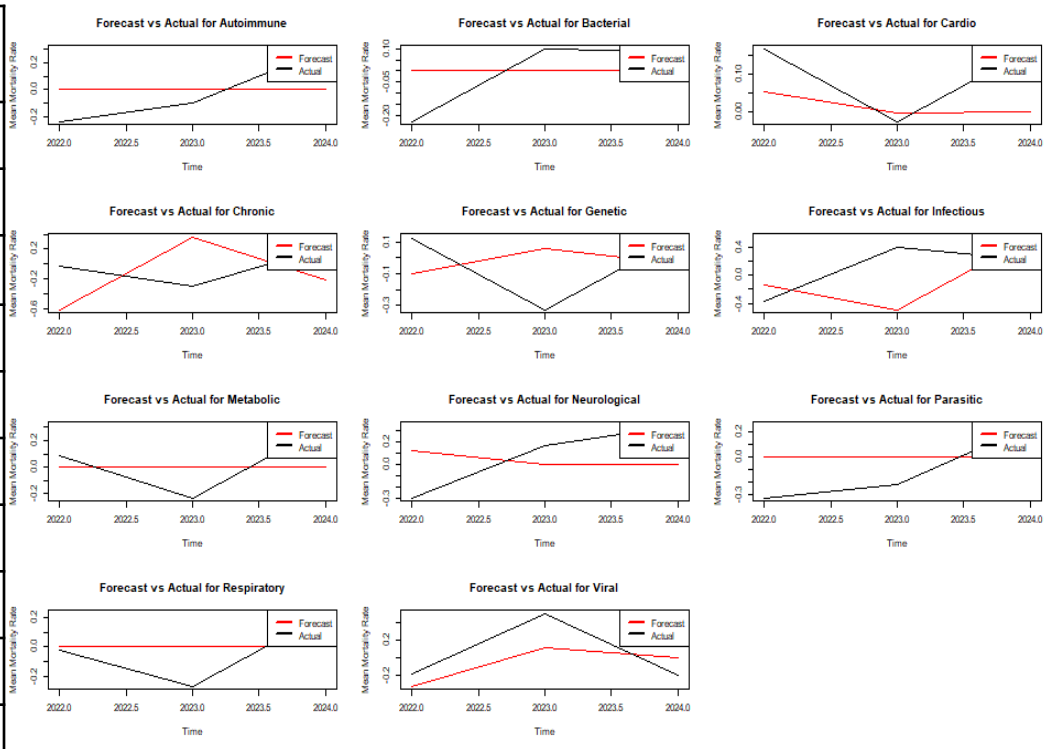
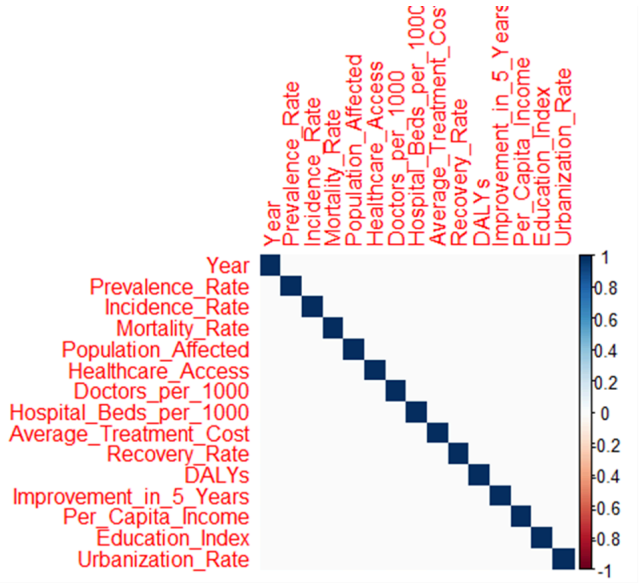


Figure 3. Plot of forecast (red line) vs actual(black line) values of the mean mortality for each disease category.

Dynamic Panel Data Model

Model Selection and Justification

This dataset contains a handful of external variables other than the mean mortality rate, which opens the possibility of using principal component analysis (PCA). One of the requirements of PCA is that multicollinearity needs to exist between the features. The correlational matrix and variance inflation factor (VIF) are calculated to investigate the correlation between features. Figure 4 shows the result that there is no correlation at all between the numeric and nonnumeric features.




```
Loading required package: carData
                                      GVIF Df GVIF^(1/(2*Df))
Country                           1.001029 19        1.000027
Year                              1.000075  1        1.000038
Disease_Name                      1.001013 19        1.000027
Disease_Category                  1.000654 10        1.000033
Prevalence_Rate                   1.000071  1        1.000036
Incidence_Rate                    1.000051  1        1.000026
Age_Group                         1.000200  3        1.000033
Gender                            1.000154  2        1.000038
Population_Affected               1.000077  1        1.000039
Healthcare_Access                 1.000076  1        1.000038
Doctors_per_1000                  1.000068  1        1.000034
Hospital_Beds_per_1000            1.000083  1        1.000042
Treatment_Type                    1.000218  3        1.000036
Average_Treatment_Cost            1.000114  1        1.000057
Availability_of_Vaccines_Treatment 1.000044  1        1.000022
Recovery_Rate                     1.000054  1        1.000027
DALYs                             1.000059  1        1.000030
Improvement_in_5_Years            1.000068  1        1.000034
Per_Capita_Income                 1.000071  1        1.000035
Education_Index                   1.000080  1        1.000040
Urbanization_Rate                 1.000078  1        1.000039
```

Figure 4. On the left side is a visualization of a correlational matrix of all numeric variables. On the right side is the output of VIF of all variables.

The next model that is suitable for this dataset, which contains categorical variables of country and disease category, is the dynamic panel data model. The dynamic panel data model allows for investigation of the ordinary linear model or generalized linear model of the cross-sectional dimensions in the panel data format, where the time series can be separated into different countries and disease categories. The stationarity of the panel dataset was tested using the Augmented Dickey-Fuller Test. The significance of the p-value is 0.01, which is lower than 5%, indicating that the series rejects the null hypothesis of the presence of a unit root and non-stationary series. Afterward, the panel data was fitted to three types of linear models, including the fixed effects model, the random two-way effects model, and the between model. A fixed effects model explains the variation in mean disease mortality within each country and disease category. The model eliminates the effect on dependent variables for any features or independent variables that are fixed or do not change with time. Examples of the variables that would be controlled by this model are a country's culture, genetics, or geography. The assumption of the fixed effects model is that unobserved traits are correlated with predictors.

The random effects model explains the variation in mean disease mortality both within and between each country and disease category. The model assumes that the unobserved traits (such as a country's culture, genetics, or geography) are not correlated with the independent variables. Therefore, the model allows us to explore the effect of differences in mean mortality rate across different countries, years, and disease categories. Lastly, the between model investigates the cross-sectional relationship, differences between countries and disease categories, while disregarding the changes over time within each unit. The linear equation for all these models is:

mean_mortality = lag(mean_mortality) + mean_prevalence + mean_income + mean_incidence + mean_healthcare + mean_doc_per1000 + mean_treatment_cost + mean_hos_per1000 + mean_treatment_cost + mean_recovery + mean_DALYs + mean_imp_5_yrs + mean_urban

The table below contains the analysis results of the three models:

Random model: within and between	Coefficients: Estimate Std. Error z-value Pr(>|z|) (Intercept) 4.3657e+00 4.5710e-01 9.5508 <2e-16 *** lag(mean_mortality) -3.5966e-03 1.6848e-02 -0.2135 0.8310 mean_prevalence -1.5011e-04 8.3236e-03 -0.0180 0.9856 mean_income 2.4402e-06 1.7009e-06 1.4347 0.1514 mean_incidence 8.0395e-03 1.1242e-02 0.7151 0.4745 mean_healthcare 3.5254e-03 3.3312e-03 1.0583 0.2899 mean_doc_per1000 -3.1555e-04 3.7178e-02 -0.0085 0.9932 mean_treatment_cost -4.6768e-06 3.3341e-06 -1.4027 0.1607 mean_hos_per1000 2.6823e-02 1.7653e-02 1.5194 0.1287 mean_recovery 8.4195e-04 3.3729e-03 0.2496 0.8029 mean_DALYs -4.8403e-05 3.3336e-05 -1.4520 0.1465 mean_imp_5_yrs 1.5427e-02 1.6384e-02 0.9416 0.3464 mean_urban 3.9059e-03 2.3752e-03 1.6445 0.1001 --- Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 Total Sum of Squares: 158.42 Residual Sum of Squares: 157.82 R-Squared: 0.0037996 Adj. R-Squared: 0.00039082 Chisq: 13.3758 on 12 DF, p-value: 0.34232
Between model	Coefficients: Estimate Std. Error t-value Pr(>|t|) (Intercept) 1.2807e-01 1.9953e+00 0.0642 0.95062 lag(mean_mortality) 7.5577e-01 7.5780e-02 9.9733 2.178e-05 *** mean_prevalence -3.4141e-02 3.5801e-02 -0.9537 0.37202 mean_income 2.3668e-05 8.2639e-06 2.8640 0.02420 * mean_incidence 5.9799e-02 5.4019e-02 1.1070 0.30488 mean_healthcare 2.2765e-02 1.4380e-02 1.5831 0.15741 mean_doc_per1000 4.6232e-02 1.5311e-01 0.3020 0.77146 mean_treatment_cost -1.2367e-06 1.0730e-05 -0.1153 0.91148 mean_hos_per1000 -1.2338e-01 9.5641e-02 -1.2901 0.23801 mean_recovery -3.8895e-02 1.3660e-02 -2.8473 0.02478 *

	mean_DALYs 3.1249e-06 1.6235e-04 0.0192 0.98518 mean_imp_5_yrs -1.2188e-02 8.5605e-02 -0.1424 0.89080 mean_urban 2.9140e-02 1.2767e-02 2.2825 0.05643 . --- Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 Total Sum of Squares: 0.0050897 Residual Sum of Squares: 9.3193e-05 R-Squared: 0.98169 Adj. R-Squared: 0.9503 F-statistic: 31.2749 on 12 and 7 DF, p-value: 6.6087e-05
Fixed effect model: within	Coefficients: Estimate Std. Error t-value Pr(>|t|) lag(mean_mortality) -9.0978e-03 1.6896e-02 -0.5385 0.5903 mean_prevalence -3.9525e-04 8.3582e-03 -0.0473 0.9623 mean_income 2.2727e-06 1.7054e-06 1.3327 0.1827 mean_incidence 5.6422e-03 1.1274e-02 0.5005 0.6168 mean_healthcare 3.5008e-03 3.3415e-03 1.0477 0.2949 mean_doc_per1000 3.9975e-04 3.7296e-02 0.0107 0.9914 mean_treatment_cost -5.1394e-06 3.3509e-06 -1.5337 0.1252 mean_hos_per1000 2.5912e-02 1.7686e-02 1.4651 0.1430 mean_recovery 1.3460e-03 3.3821e-03 0.3980 0.6907 mean_DALYs -4.9326e-05 3.3400e-05 -1.4768 0.1398 mean_imp_5_yrs 1.5056e-02 1.6410e-02 0.9175 0.3590 mean_urban 3.5975e-03 2.3815e-03 1.5106 0.1310 Total Sum of Squares: 157.53 Residual Sum of Squares: 156.94 R-Squared: 0.0037225 Adj. R-Squared: -0.005132 F-statistic: 1.08605 on 12 and 3488 DF, p-value: 0.36709

For the random model, only the intercept is a significant factor, and no other independent variables are important or explain the mean mortality rate. The R-squared is also low, suggesting that the model only explains 0.38% of the variance in mean mortality within the group. The chi-square p-value of the random model is also higher than 5%, which suggests the model does not improve when compared to a null model. For the between model, the first lag of mean mortality rate, mean income, mean urbanization rate, and mean recovery rate are statistically significant in explaining the difference between the country and disease group. The R-squared is high with 98.2%, which means the model explains 98.2% of the variance in mean mortality between groups. The F-statistic p-value is very low and below 5%, suggesting that the model is statistically improved from the null model. For the within model, there is no significant factor, and no other independent variables are important or explain the mean mortality rate. The R-squared is also low, suggesting that the model only explains 0.37% of the variance in mean mortality within and between the groups. The F-statistic p-value of the random model is also higher than 5%, which suggests the model does not improve when compared to a null model.

Out-of-sample Validation

The next step after fitting the panel data models to the time series is to forecast the mean mortality rate. The RMSE of the predicted value from models is calculated as shown in the table below. The train dataset is from 2001 to 2017, and the test dataset is from the year 2018 to the year 2024. Between models perform the best in predicting the mean mortality, while Random and Fixed effects are equal in their performance in predicting out of sample. The visualization of the prediction of three models versus the actual model is shown in Figure 5. Panel data from Japan was used as an example here. From the graph, we can see that the Between model prediction line has a fluctuation that is a little delayed from the actual data. On the other hand, random and fixed effect predictions appear as almost a straight line or plateau with minor fluctuation.

Models	RMSE values
Between	0.140
Random	0.22
Fixed Effects	0.22

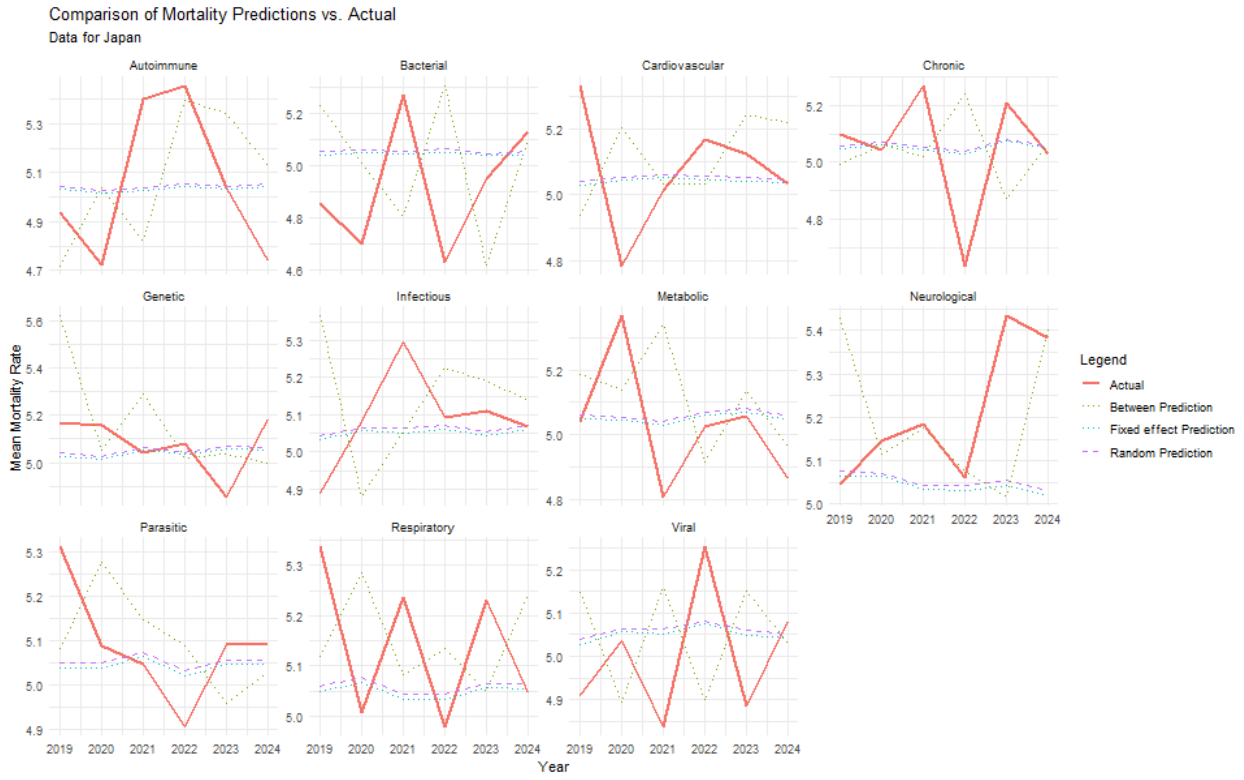


Figure 5. Facet grid plots of the comparison between predictions vs actual mean mortality in each disease category of Japanese dataset.

Narrative and Insights

Major Outbreak of the 2000s

The Human race faces a new era in the 2000s with the West Nile Virus, which is transmitted by mosquito bites (Moran-Perez, 2025). The countries with warmer temperatures are more prone to outbreaks. The duration of the West Nile Virus outbreak lasted three years. In that duration, there were 4,156 cases and 284 deaths in America. After that, humans faced a bigger viral respiratory illness globally, which was SARS-CoV beginning in 2003 and MERS-CoV in 2012. SARS-CoV caused the death of 774 people and 858 deaths (Mayo Clinic, 2021). Even though humans have persevered through many viral illnesses at this point, Coronavirus-19 will be the biggest outbreak of the 21st century. COVID-19 is similar to SARS and MERS in that they belong to a larger family of coronaviruses that cause upper-respiratory tract illness in humans (National Institute of Allergy and Infectious Diseases, 2024). COVID-19 is different from SARS and MERS because they are more infectious and virulent than the other two. The total number of deaths is approximately seven million from COVID-19 (Worldometer, 2024).

Story of United States Disease's Mortality

Based on the history of the mean mortality of eleven disease categories in the United States, we have seen that the mortality rates are always fluctuating around the mean, and there are no consistent trends across all the categories. Viral disease category is the one with a clear upward trend throughout the time, which is due to the COVID-19 pandemic in recent years and SARS-CoV and MERS-CoV in the early 200s. By applying the ARIMA model to the timeseries, we have learned that mortality from autoimmune, bacterial, metabolic, parasitic, and respiratory disease types cannot be determined from only their past values and shocks. On the other hand, chronic, genetic disease, and infectious disease mean mortality rate can be modeled by past values, and cardiovascular, neurological, and viral diseases can be modeled by past shocks. Based on the out-of-sample validation, the disease category that can be modeled and has the best prediction of future mean mortality rate is Cardiovascular disease.

Story of Global Disease's Mortality

We have explored the past history of disease's mean mortality in the twenty countries and found that not every country experiences the same trends. In the 2000s, there were many major outbreaks of respiratory disease viruses, but not every country's mortality rate of viral or respiratory diseases was affected by this. Therefore, we are investigating whether external factors have any influence on the differences in mortality rate among twenty countries by using dynamic panel data on fixed effect, between, and random models. There are no significant effects of the external factors and past lag of the data itself on the mean mortality rate in fixed-effect and random models. On the other hand, the first lag of mean mortality rate, mean income, mean urbanization rate, and mean recovery rate of the between models are statistically significant on the mean mortality rate. Moreover, the between model is also the best performer in predicting the

mortality rate. This tells us that the factor that cross-sectional variation impacts the mean mortality from the disease in a country more than the time-based variation.

Citations

Brigham, Kenneth L. “Predictive Health: The Imminent Revolution in Health Care.” *Journal of the American Geriatrics Society*, vol. 58, no. s2, Oct. 2010, pmc.ncbi.nlm.nih.gov/articles/PMC3760012/#R1, https://doi.org/10.1111/j.1532-5415.2010.03107.x. Accessed 22 Apr. 2025.

Mayo Clinic. “History of Infectious Disease Outbreaks and Vaccines Timeline.” *Mayo Clinic*, 2021, www.mayoclinic.org/diseases-conditions/history-disease-outbreaks-vaccine-timeline. Accessed 12 May 2025.

MalaiarasuGRaj. (2024). Global Health Statistics [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/10028650

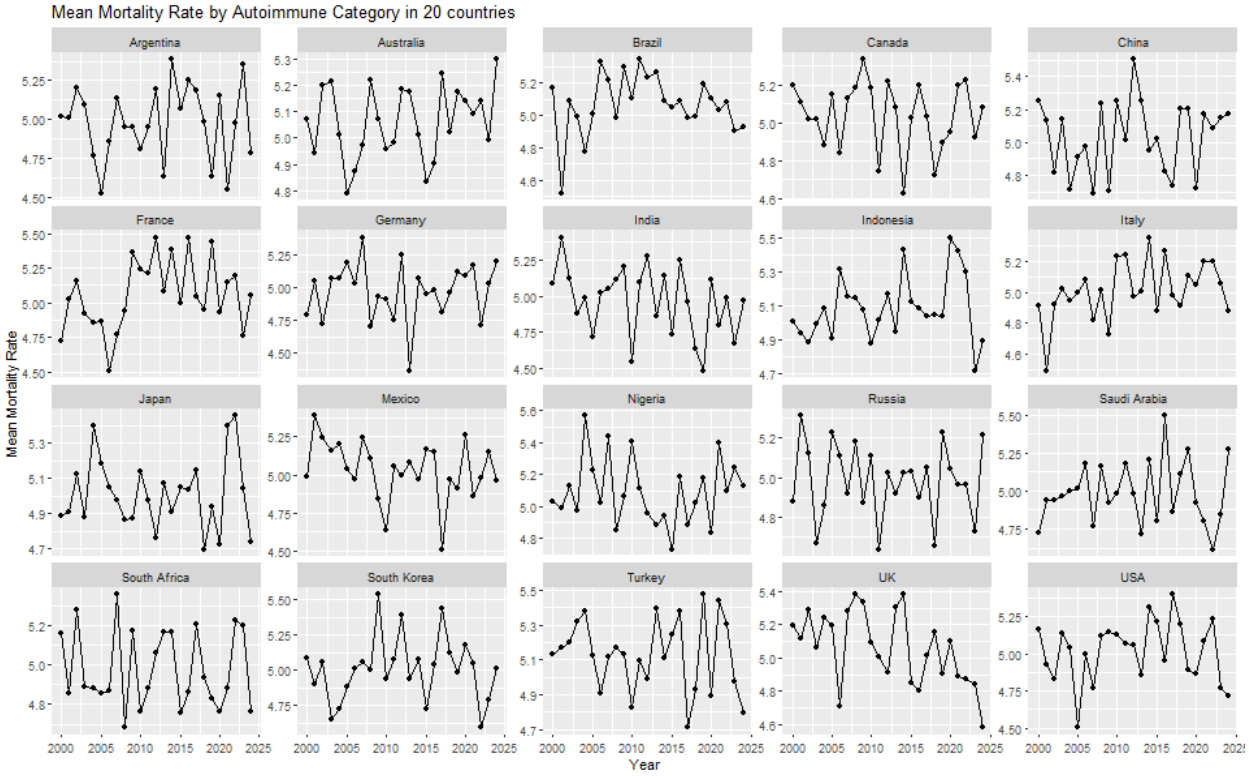
Moran-Perez, Gillian. “A Masterclass by Laura Brenes.” *Daily Sundial*, 9 May 2025, sundial.csun.edu/156361/news/a-timeline-of-outbreaks-from-2000-to-present/. Accessed 12 May 2025.

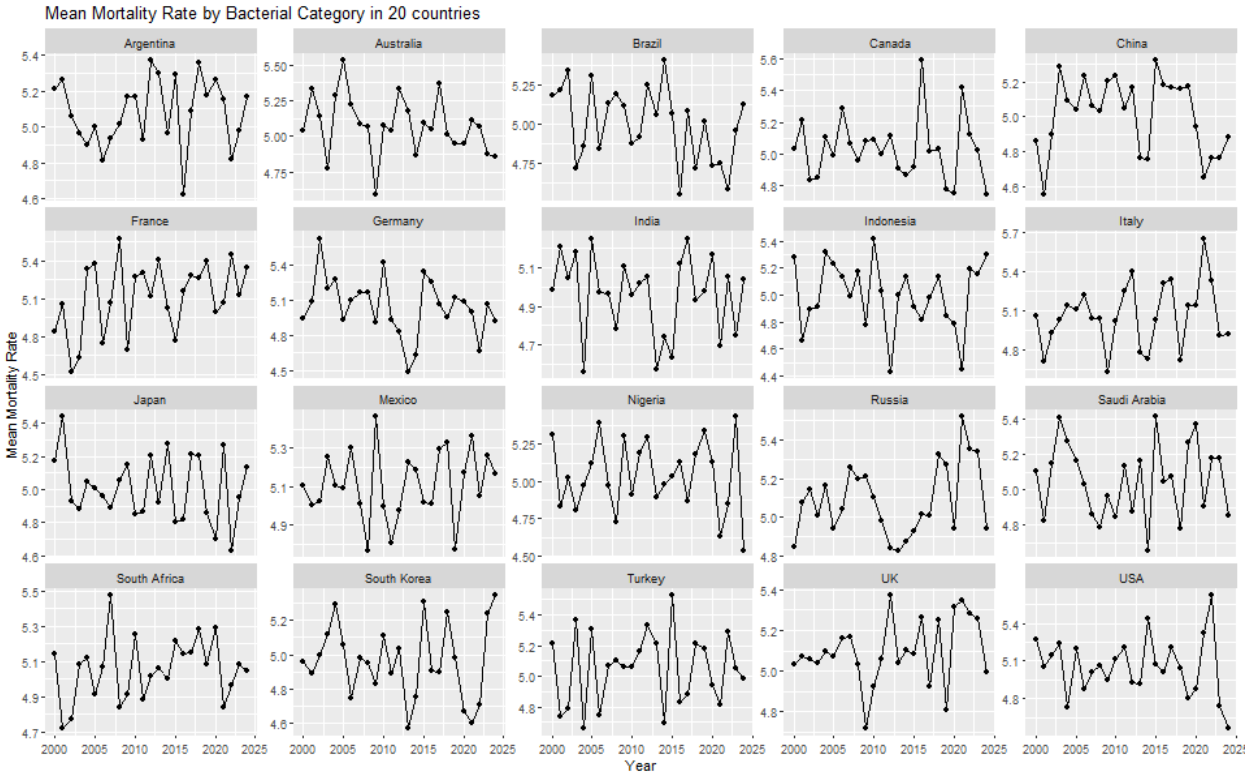
National Institute of Allergy and Infectious Diseases. “Coronaviruses.” *Nih.gov*, 31 July 2024, www.niaid.nih.gov/diseases-conditions/coronaviruses. Accessed 12 May 2025.

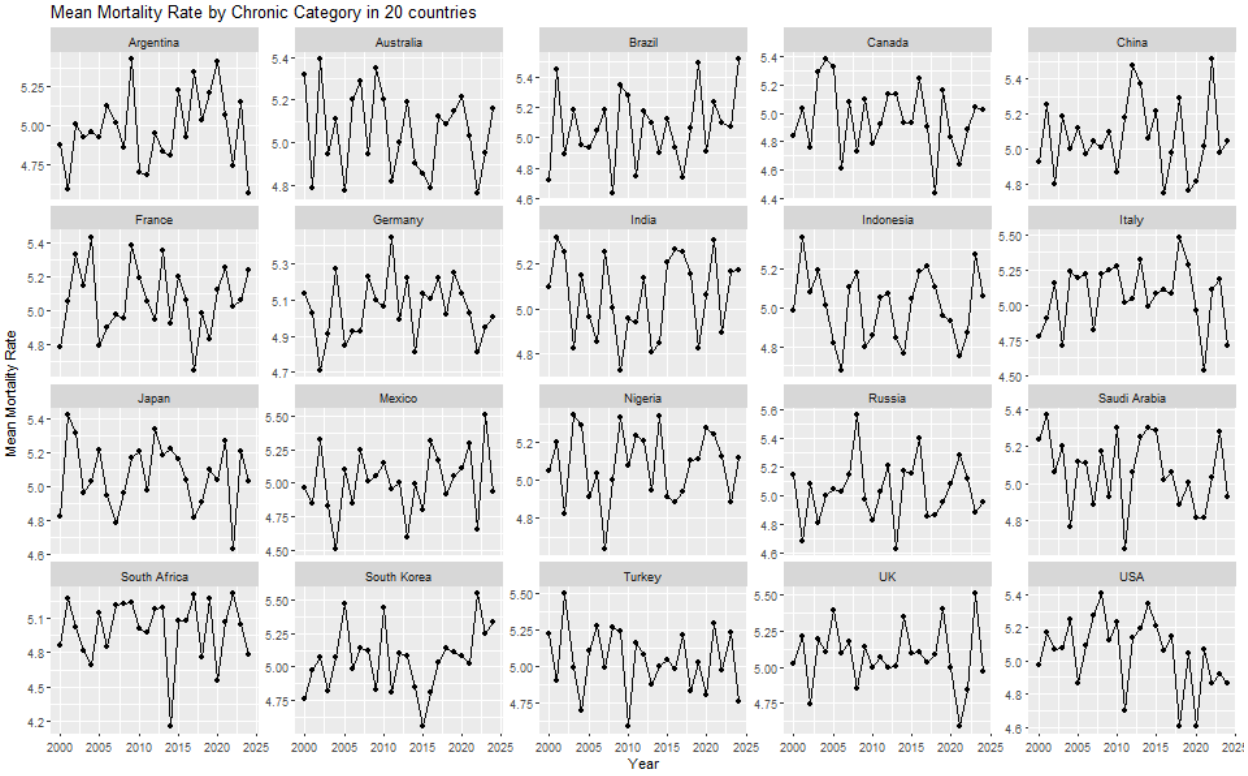
Nolte E, McKee CM. Measuring the health of nations: Updating an earlier analysis. Health Aff. 2008;27:58–71. doi: 10.1377/hlthaff.27.1.58.

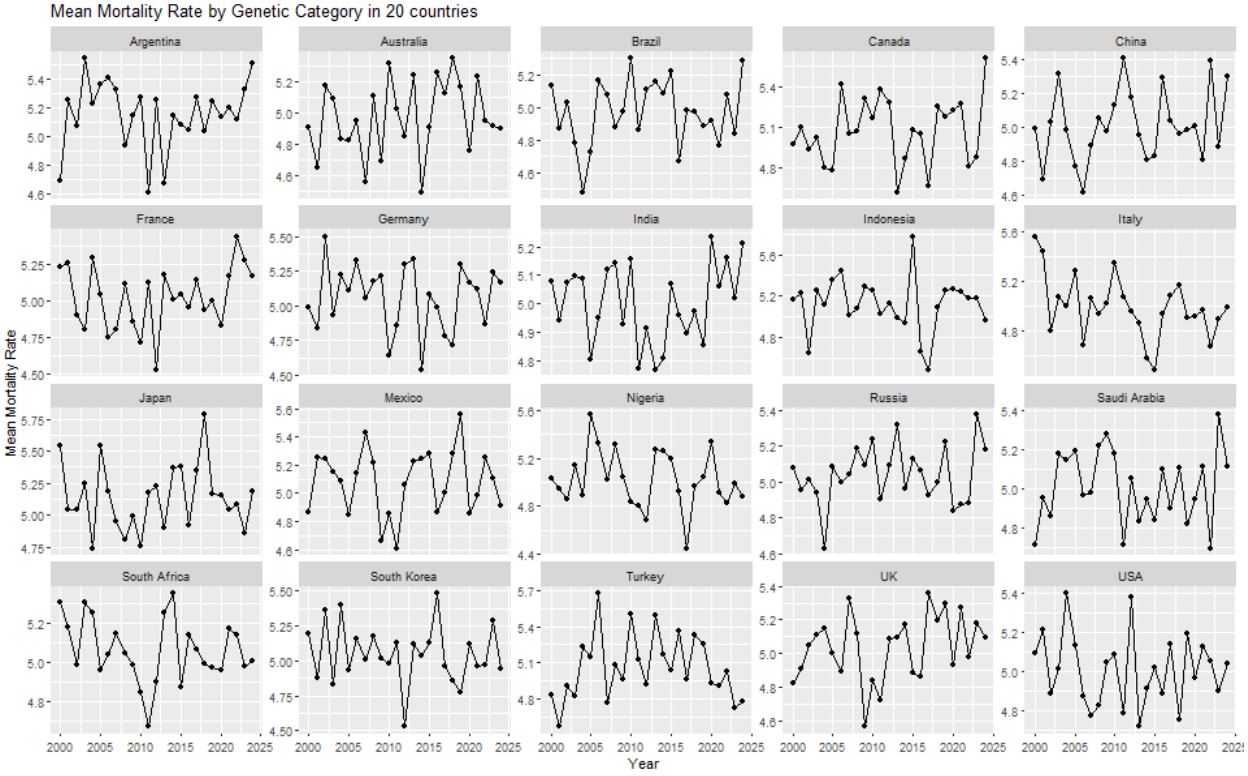
Worldometer. “COVID - Coronavirus Statistics - Worldometer.” *Worldometers.info*, 2024, www.worldometers.info/coronavirus/. Accessed 12 May 2025.

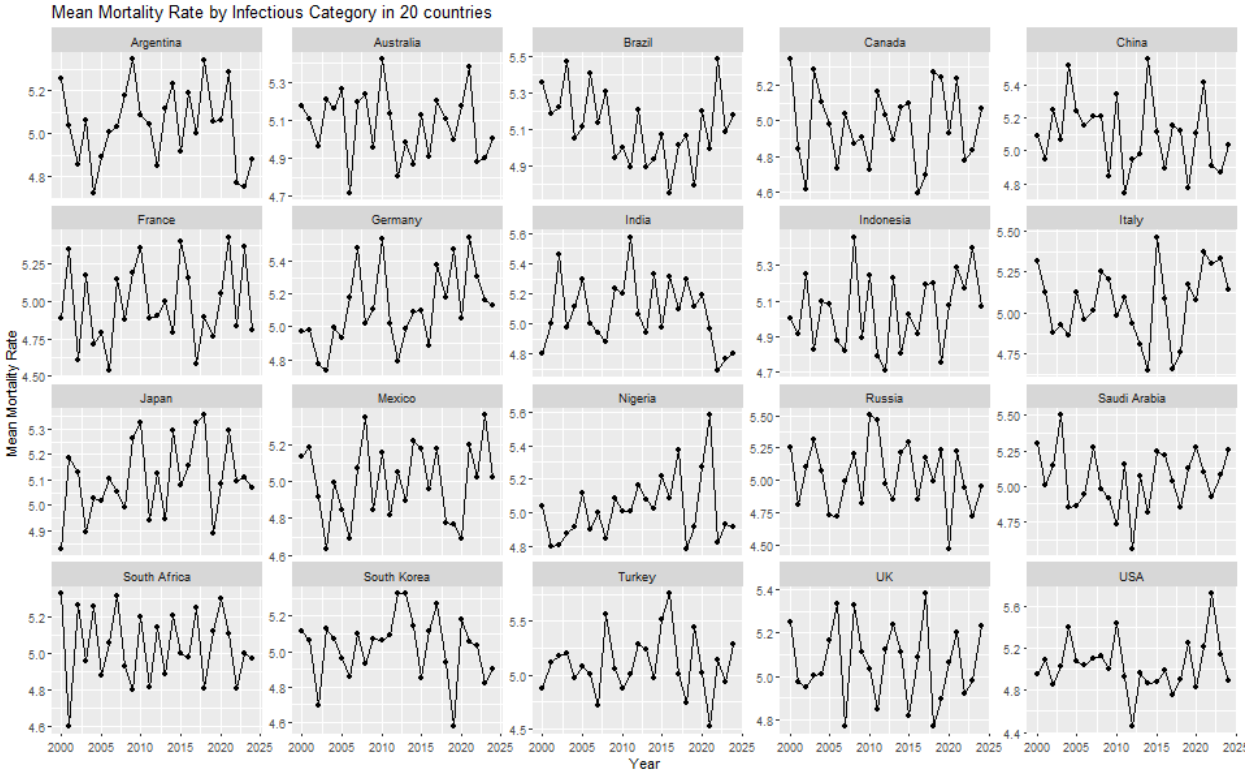
Appendix

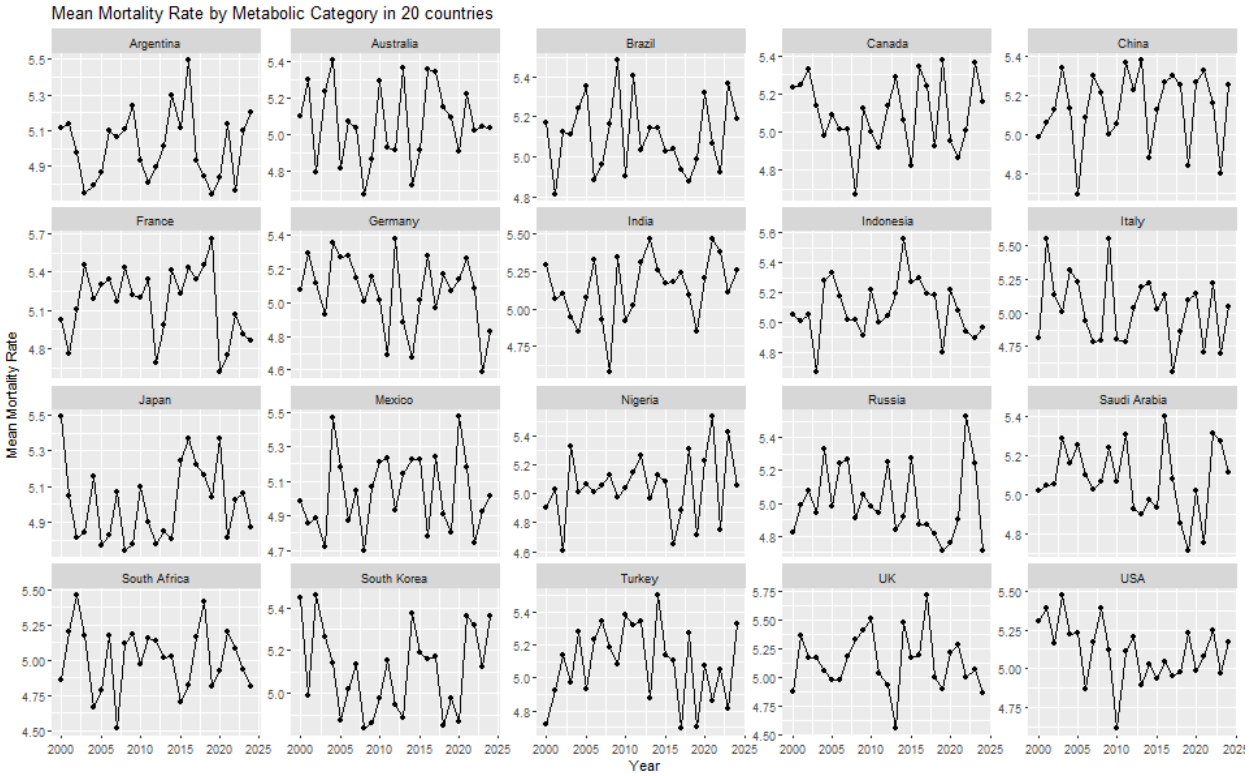

Mean Mortality Rate by Autoimmune Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year

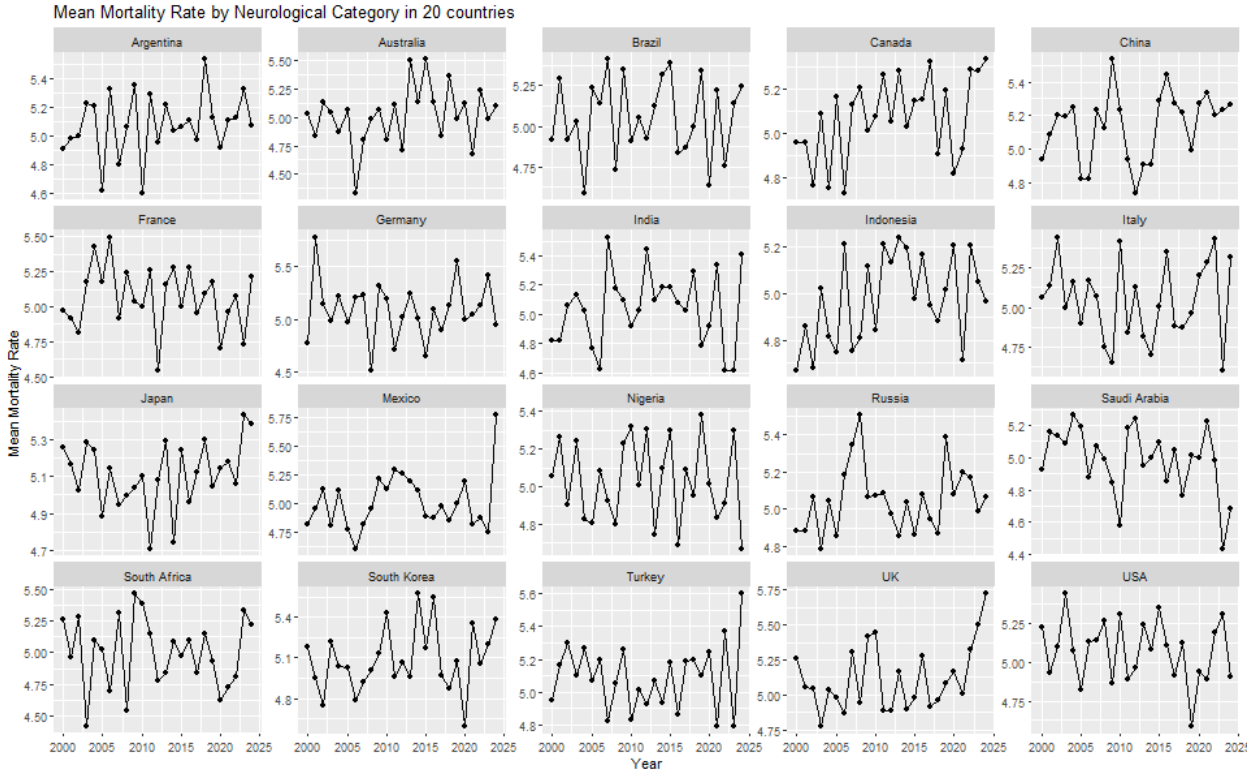

Mean Mortality Rate by Bacterial Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year

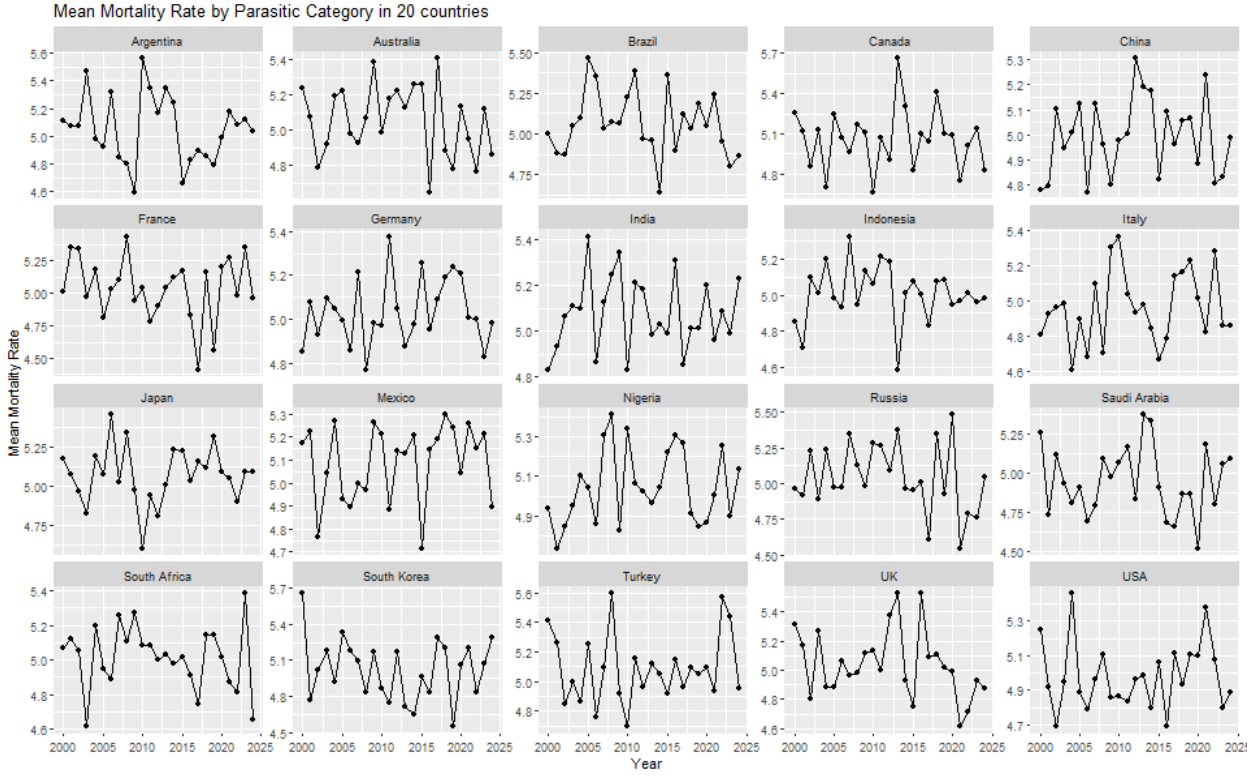

Mean Mortality Rate by Chronic Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year

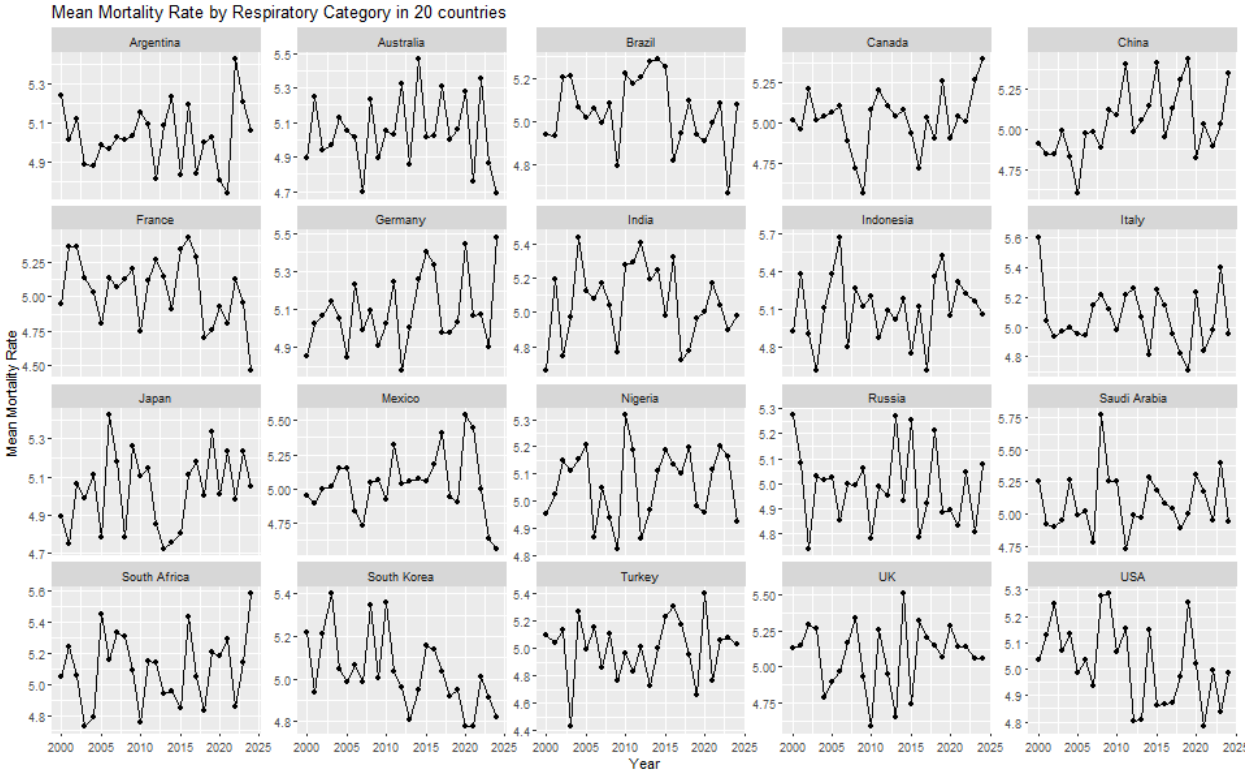

Mean Mortality Rate by Genetic Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year


Mean Mortality Rate by Infectious Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year


Mean Mortality Rate by Metabolic Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year


Mean Mortality Rate by Neurological Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year


Mean Mortality Rate by Parasitic Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year


Mean Mortality Rate by Respiratory Category in 20 countries
Argentina
Australia
Brazil
Canada
China
France
Germany
India
Indonesia
Italy
Japan
Mexico
Nigeria
Russia
Saudi Arabia
South Africa
South Korea
Turkey
UK
USA
Mean Mortality Rate
Year