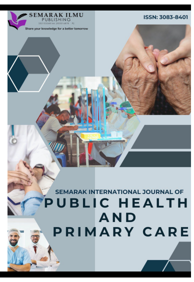
Semarak International Journal of Public Health and Primary Care

Journal homepage: https://semarakilmu.my/index.php/sijphpc/index
ISSN: 3083-8401



The Effects of Non-Communicable Diseases and their Risk Factors on Malaysia's Healthcare Expenditures

Fatin Nadia Zulfakairuddin[1], Sharifah Fairuz Syed Mohamad[1,*], Mazlynda Md. Yusuf[1], Nurul Sima Mohamad Shariff[1] , Sulieman Yhia[2]

[1] Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800 Nilai, Negeri Sembilan, Malaysia
[2] The Gulf international Private Academy, Al-Ain, United Arab Emirates

ARTICLE INFO

Article history:
Received 2 September 2025
Received in revised form 23 September 2025
Accepted 5 October 2025
Available online 16 October 2025

Keywords:
Cardiovascular Disease (CVD); diabetes; dietary risk; healthcare expenditure; High Body-Mass Index (BMI); Non-Communicable Diseases (NCDs); smoking

ABSTRACT

Malaysia's healthcare expenditures are heavily burdened by a rise in the rate of non-communicable diseases (NCDs), specifically diabetes and cardiovascular disease (CVD). This research examines how these non-communicable diseases (NCDs) are impacted by selected key risk factors, which are smoking, dietary risk, and having a high body-mass index (BMI), as well as how these NCDs affect Malaysia's healthcare expenditure. Using a 20-year dataset (2000–2021) from the World Bank and Global Health Data Exchange (GHDX), the study uses time series analysis and multiple linear regression. Results indicate that high BMI has the strongest correlation with both diseases. In addition, diabetes is highly correlated with healthcare expenditure in Malaysia. These results highlight the need of focused prevention efforts that address modifiable risk factors in order to decrease the financial burden of NCDs. Hence, to reduce the occurrence of these diseases, policymakers are recommended to give priority to providing funds for efficient management plans and public awareness initiatives. A comprehensive approach to addressing these risk factors may enhance public health outcomes and lower Malaysian healthcare expenditures.

1. Introduction

According to Stanford University [20], risk can be defined as the probability that an incident might negatively impact the achievement of the organization's objectives. This concept emphasizes how crucial it is to recognize and control risks in order to protect and improve organizational performance. In other words, risk might affect many types of organizations such as healthcare. Therefore, healthcare risk is classified as an event that could impact the delivery of healthcare [20]. This explains that healthcare risk are the circumstances or events that related to patient care and safety. As a result, risk management for healthcare risks is essential. According to Catalyst [3], in healthcare risk management, risks are identified, tracked, evaluated, reduced, and eliminated via the use of clinical

* *Corresponding author.*
E-mail address: fatinnadiazulfakairuddin@gmail.com

https://doi.org/10.37934.sijphpc.6.1.2743

and administrative systems, procedures, and reports. As a result, healthcare risk management will promote patient safety the provision of healthcare as a whole.

Besides, non-communicable diseases (NCDs) are characterized as diseases that are typically caused by unhealthy behaviors instead of infections or social contact according to International Federation of Red Cross and Red Crescent Societies [11]. This shows that rather than being spread from person to person, these diseases arise because of the patient's own risk factors. According to World Health Organization [29], the primary causes of noncommunicable diseases are three specific behaviors which are using tobacco products, not exercise and eating poorly. Additionally, non-communicable diseases examples include cancer, diabetes, hypertension, and cardiovascular disorders.

In Malaysia, non-communicable diseases rank among the main causes of death according to Worl Health Organization [27]. According to World Health Organization [29], the most prevalent NCDs are chronic respiratory conditions, cancer, diabetes, heart disease, and stroke. According to Diana *et al.*, [8], as risk factors like poor diet, inactivity, alcohol and tobacco use, and metabolic risk factors like high body-mass index, diabetes and cardiovascular diseases worsen, there will probably be a rise in NCD-related mortality in the coming decades. According to Chandran *et al.*, [7], over 70% of the disease burden and 67% of premature deaths in Malaysia in 2014 were caused by NCDs. This explains that NCDs take over more than half of the cause of death in Malaysia.

Additionally, diabetes is a chronic illness that arises from inability of the body to effectively use or lack of production of insulin by the pancreas and the hormone insulin is one that regulates blood sugar [28]. This explains that when the body cannot utilize the insulin properly, it will lead to higher blood sugar levels that can cause severe health condition. According to World Health Organization [28], diabetes is mainly categorized into two types which are type 1 whereby the pancreas does not create much insulin on its own and type 2 which this happens when the body doesn't produce enough insulin or gets resistant to it.

In addition, cardiovascular disease is a medical issue that affects the heart or blood vessels World Health Organization [26]. According to National Health Service UK (n.d.), In addition to a higher risk of blood clots, cardiovascular disease is commonly associated with atherosclerosis, or the build-up of fat inside the blood vessel and it can also be linked to arterial damage in the kidneys, eyes, heart, and brain. The four types of cardiovascular diseases include strokes, coronary heart disease, peripheral arterial disease, and aortic disease National Health Service UK [17]. As a result, cardiovascular disease is among the primary reasons of death worldwide [26]. This emphasizes that cardiovascular diseases represent a danger to global health.

Furthermore, in Malaysia, most medical costs were linked to non-communicable diseases. This is because long-term care and expensive medications are necessary for the treatment and maintenance of NCDs resulting in a significant financial burden on the healthcare system and the people receiving it. According to Ministry of Health Malaysia [14], the total costs for hospitalization for certain NCDs which are cardiovascular disease, diabetes and cancer are RM 2,123,786,024. This explains that the growing number of patients with NCDs is causing hospitals to become overcrowded and waiting times to lengthen, which is impacting the system's overall effectiveness.

Moreover, according to Lim [12], the incapacity of Malaysians to apply knowledge to prevent non-communicable diseases is the factor causing NCDs in Malaysia. There is a difference between being aware of the risk factors of NCDs and really take steps in preventing them such as having a healthier diet, exercising and decreasing the usage of tobacco and alcohol. In Malaysia, sedentary lifestyles and poor eating habits are frequently encouraged by culture, which improves the country's increasing NCD rate such as a large intake of fatty and sugary meals. Additionally, mental health conditions like stress, and anxiety make it harder to sustain healthy practices. Therefore, given the previous discussion, it implies that understanding NCDs and its risk factors is vital, especially how it affects Malaysia's healthcare expenditure.

Using a quantitative methodology, this study examines the relationship between non-communicable diseases (NCDs) and their impact on healthcare spending in Malaysia. In order to analyze the secondary data that obtained from the World Bank and Global Health Data Exchange (GHDX), this study used time series analysis. Multiple linear regression is the methods of statistical analysis used. This research objectives are integral to the study's framework:

1) To determine which risk factors (smoking, high body-mass index (BMI) and dietary risk) will have the most effects on each of the non-communicable diseases (NCDs) under study, namely diabetes and cardiovascular diseases.
2) To examine which of the non-communicable diseases (NCDs) contribute for the most portion of Malaysia's healthcare expenditures.

In order to achieve these objectives, this research will answer the following questions:

1) Which of the risk factors (smoking, high body-mass index (BMI) and dietary risk) will have the most effects on each of the non-communicable diseases (NCDs) under study, namely diabetes and cardiovascular diseases?
2) Which of the non-communicable diseases (NCDs) contribute for the most portion of Malaysia's healthcare expenditures?

2. Literature Review

2.1 Risk Factors and Non-Communicable Diseases (NCDs)

Diabetes and cardiovascular diseases are among the non-communicable diseases (NCDs) that contribute most to Malaysia's health issues. A high body-mass index (BMI), smoking, and poor diet have been identified as important risk factors for these diseases' onset.

Many NCDs are known to be significantly increased by smoking. according to Lushniak *et al.,* [13], smoking, including second-hand smoke, has been connected to diabetes, cardiovascular issues, and chronic respiratory disorders. Stronger tobacco control laws are therefore imperative. Smoking and noncommunicable diseases (NCDs) are known to be linked, but it is crucial to understand that smoking may not be the only factor contributing to the rising number of NCD cases. Other factors, such as the environment, access to treatment, and socioeconomic status, are also contributing to the rise in these diseases.

Another major risk factor for the diseases such as diabetes and cardiovascular disease is having a high body-mass index (BMI) [29]. Obesity increases the risk of conditions of these diseases. The significance of maintaining a healthy weight is emphasized by the World Cancer Research Fund International [25], which also supports this relation because BMI does not account for the distribution of fat in the body or differentiate between muscle and fat, it is crucial to remember that it is not a perfect indicator of health.

Another major cause of NCDs is dietary risk, specifically a diet heavy in processed foods and low in fruits and vegetables. Insulin resistance, high blood pressure, and high cholesterol are all linked to diabetes and cardiovascular disease and can be increased by poor diet according to Centers for Disease Control and Prevention [6].

As stated by the World Cancer Research Fund International [25], plant-based diets may reduce the chance of developing certain illnesses. However, it is crucial to consider that dietary recommendations and guideline might not benefit equally to all groups since some of it have a limited access. Thus, it is necessary to focus on the food accessibility and to control the portion based on the food pyramid.

H1: Dietary risk (DIETR) has a significant positive impact on cardiovascular disease (CVD).
H2: High body-mass index (HBMI) has a significant positive impact on cardiovascular disease (CVD).
H3: Smoking (SM) has a significant positive impact on cardiovascular disease (CVD).
H4: Dietary risk (DIETR) has a significant positive impact on diabetes (DIAB).
H5: High body-mass index (HBMI) has a significant positive impact on diabetes (DIAB).
H6: Smoking (SM) has a significant positive impact on diabetes (DIAB).

2.2 Non-Communicable Diseases (NCDs) and Healthcare Expenditure

Healthcare systems are heavily impacted by a rising prevalence of diabetes and cardiovascular disease (CVD). In 2022, the direct medical expenses for noncommunicable diseases (NCDs), especially diabetes, amounted to RM 4.38 billion in Malaysia, representing 45.38% of the total expenditures of NCDs according to Ministry of Health Malaysia [16]. This emphasizes how important it is to have effective prevention and treatment plans.

Diabetes has a significant financial impact on healthcare expenses according to Sun *et al.,* [21]. However, their research did not completely take into consideration factors that also contribute to increased costs, such as medicine costs, administrative costs, and the healthcare system. Similarly, the Ministry of Health Malaysia [15] has identified CVD as an economic concern due to the increasing expense of treatment.

To conclude, diabetes and cardiovascular disease have a significant financial effect on healthcare systems due to their significant direct medical expenses. Current research emphasizes their economic impact, but it frequently ignores other aspects, such as administrative and pharmaceutical expenses. In order to provide affordable healthcare solutions, a better strategy needed for better prevention and management.

H7: Cardiovascular disease (CVD) has a significant positive impact on healthcare expenditure (HEX).
H8: Diabetes (DIAB) has a significant positive impact on healthcare expenditure (HEX).

3. Methodology

3.1 Data Collection

This study focuses on two non-communicable diseases (NCDs) in Malaysia which are diabetes and cardiovascular disease and their associated risk factors which are smoking, dietary risks, and high body mass index (BMI). The data was collected over a 20-year period, from 2000 to 2021. The data on healthcare expenditures, the dependent variable, was sourced from the World Bank, while data on the NCDs and their risk factors was obtained from the Global Health Data Exchange (GHDX).

By examining these variables, this research aims to explore how Malaysia’s healthcare spending relates to these diseases and risk factors. Table 1 provides a summary of the variables, their abbreviations, and their descriptions.

Table 1
Variables, abbreviation and description on this study

Variable	Abbreviation	Description
Healthcare Expenditure	HEX	Malaysia's healthcare spending, measured as current health expenditure per capita in current US dollars.
Smoking	SM	Prevalence of smoking in Malaysia, measured in DALYs (disability-adjusted life years) as a percentage of the population across all ages and genders.
High Body-Mass Index	HBMI	Prevalence of high body-mass index (BMI) in Malaysia, measured in DALYs (disability-adjusted life years) as a percentage of the population across all ages and genders.
Dietary Risk	DIETR	Dietary risk factors in Malaysia, such as unhealthy eating habits, measured in DALYs (disability-adjusted life years) as a percentage of the population across all ages and genders.

3.2 Research Model

The goal of this research is to indicate the risk factors for the NCDs which are hypertension and diabetes that will be the focus of this study in Malaysian individuals. Three models serve as the foundation for this research, which will produce a three-structured conceptual framework that supports the objectives of this research. With the first and second models addressing the first objective which is to determine which risk factors have the most significant effects on the non-communicable diseases (NCDs) under study. First model is specifically to determine the effects on cardiovascular disease meanwhile the second model are to determine the effects on diabetes. The second objective, which is to determine which of these NCDs accounts for the largest portion of Malaysia's healthcare spending, is in line with the third model. Figure 1, 2 and 3 below shows the research framework for each of the models, which contains the dependent and independent variables.

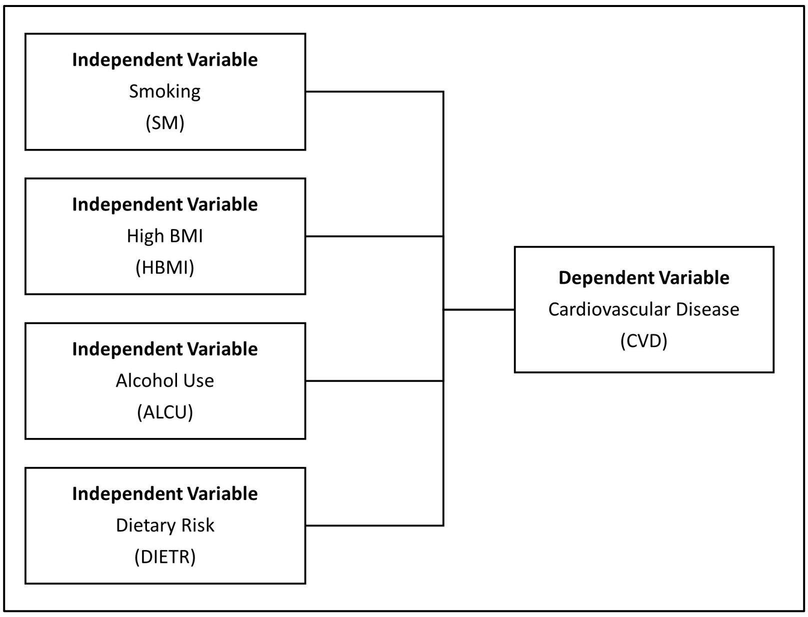


Fig. 1. Research framework for Model 1

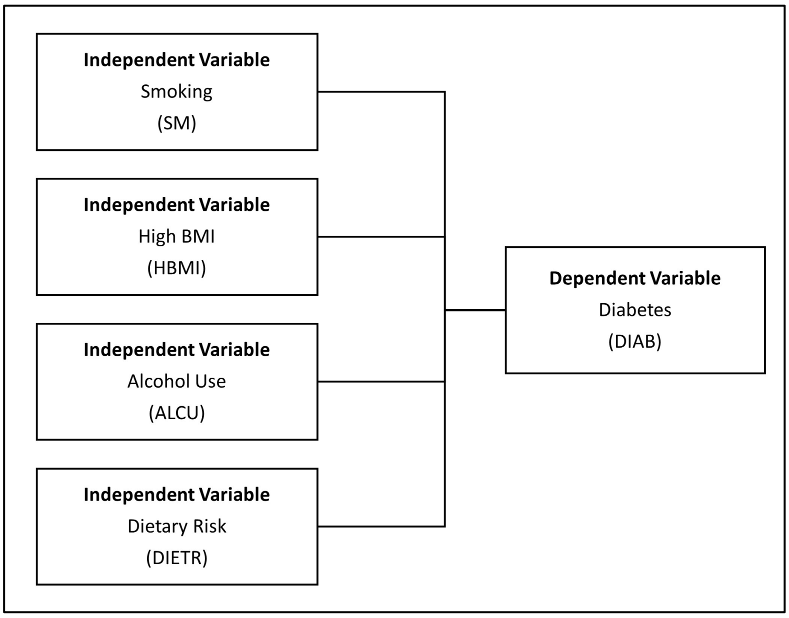


Fig. 2. Research framework for model 2

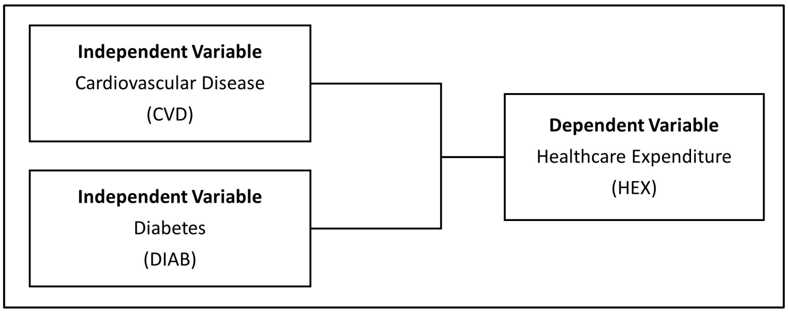


Fig. 3. Research framework for model 3

3.3 Data Analysis
3.3.1 Descriptive statistics

The data was summarized using descriptive statistics, which provided an overview of the risk factors and variables associated with non-communicable diseases (NCDs). This analysis includes calculating the mean, median and standard deviation to figure out each of the model's central tendency and variability. This initial analysis gives us a clear understanding of the dataset, setting the stage for more detailed exploration and further statistical and econometric testing.

3.3.2 Correlation analysis

Relationships between risk factors and each of the non-communicable diseases (NCDs) were examined using correlation analysis to find the association between each independent and dependent variable for each model and the relationship between each of the NCDs in this research with healthcare expenditures in Malaysia. To determine the strength of a correlation's value is when the absolute value is higher. For instance, a perfect linear relationship is shown by a correlation coefficient of -1 or +1. The weaker the relationship, the closer the correlation coefficient is to zero and no association is indicated by a value of 0.

3.3.3 Multicollinearity analysis

According to The Investopedia Team [22], when a few of independent variables in a regression model have a strong correlation with one another is referred as multicollinearity. Finding the variance inflation factor (VIF) for each independent variable is one way to find multicollinearity and if the VIF value is larger than 5 indicates multicollinearity and value that are larger than 10 indicates high multicollinearity. The correlation strength between the independent variables will be determined through VIF.

3.3.4 Augmented Dickey-Fuller (ADF)

Augmented Dickey-Fuller (ADF) test is used to figure out if a time series has a unit root or is stationary. The null hypothesis will be rejected, and the series will be considered stationary if the ADF test value is more negative than the critical value, which is at the 1%, 5%, and 10% significance level. Meanwhile, if the ADF test value is closer to 0 and less negative, the series has a unit root and is non-stationary and fail to reject the null hypothesis. Augmented Dickey-Fuller (ADF) test equation, or Eq. (1), is as follows:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-i} + \epsilon_t \quad (1)$$

where the variable being tested, Y_t, the first difference of Y_t, $\Delta Y_t = Y_t - Y_{t-1}$, constant term, α, time trend, βt, coefficient of the lagged level of Y_t, γ, lagged differences to account for autocorrelation, $\sum_{i=1}^{p} \delta_i \Delta Y_{t-i}$, error term, ϵ_t respectively.

3.3.4 Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) was used in this research to indicate the association of the independent variable and dependent variable for three model in this study. For model 1 and 2, it indicates the relationship between the associated risk factors which are smoking, high body-mass index and dietary risk and each of the non-communicable diseases (NCDs) which are cardiovascular disease for model 1 and diabetes for model 2 meanwhile model 3 is to determine the relationship between the NCDs and healthcare expenditures in Malaysia. In line with the study goals of identifying important determinants and contributors to healthcare costs, this method offers a strong framework for measuring the impact of important risk variables on illness prevalence and the ensuing economic burden. To determine which one of the independent variables are highly associated with the dependent variable, the coefficient value must be positive and the probability or the p-value must be lower than 0.05. The lowest p-value that are the nearest to 0 will indicates the most significant independent variable to the dependent variable of each of the model. Eq. (2) which is the equation for multiple linear regression (MLR) for model 1 which indicates the association between cardiovascular disease and the associated risk factors are as follows:

$$CVD = \beta_0 + \beta_1 SM + \beta_2 HBMI + \beta_3 DIETR + \varepsilon \quad (2)$$

where the variable cardiovascular disease, CVD, beta coefficient of the model (i = 1, 2, 3, … ,4), β_i, smoking, SM, high body-mass index, HBMI, dietary risk, DIETR, error term, ε, respectively.

Eq. (3) which is the multiple linear regression (MLR) for model 2, which shows the connection between diabetes and the risk variables are as follows:

$$DIAB = \beta_0 + \beta_1 SM + \beta_2 HBMI + \beta_3 DIETR + \varepsilon \quad (3)$$

where the variable diabetes, DIAB, beta coefficient of the model (i = 1, 2, 3, … ,4), β_i, smoking, SM, high body-mass index, HBMI, dietary risk, DIETR, error term, ε, respectively.

Eq. (3) which is the multiple linear regression (MLR) for model 2, which shows the connection between diabetes and the risk variables are as follows:

$$HEX = \beta_0 + \beta_1 CVD + \beta_2 DIAB + \varepsilon \quad (4)$$

where the variable healthcare expenditure, HEX, beta coefficient of the model (i = 1, 2, 3, … ,4), β_i, cardiovascular disease, CVD, diabetes, DIAB, error term, ε, respectively.

4. Results and Discussions

4.1 Stationary Test: Augmented Dickey-Fuller (ADF)

Table 2 summarizes the results of Augmented Dickey-Fuller (ADF) test which is used to determine the stationary of time series variables. In this research, the test was conducted for six variables which are CVD, DIAB, HEX, DIETR, HBMI and SM at their level, first difference and also second difference. Whether the null hypothesis of non-stationarity which determines the existence of a unit root may be rejected is determined by comparing the test's p-values to a significance level of 0.05.

Table 2
Unit root test: Augmented Dickey-Fuller (ADF)

Variable	Level	1st difference	2nd difference
Cardiovascular disease (CVD)	0.3162	0.4867	0.0137
Diabetes (DIAB)	0.9763	0.7476	0.0000
Healthcare Expenditure (HEX)	0.9959	0.0109	0.0000
Dietary Risk (DIETR)	0.6782	0.1733	0.0036
High Body-Mass Index (HBMI)	0.9998	0.1960	0.0000
Smoking (SM)	0.2353	0.1954	0.0005

At the original level, none of the variables are stationary, as all their p-values are above 0.05. This means the data at this stage likely contains trends or other patterns that make it non-stationary. For example, HEX has a very high p-value of 0.9959, clearly indicating non-stationarity. This indicates that before doing any additional analysis, the data must be transformed.

The narrative slightly shifts after making the first adjustment. HEX becomes stable, indicating that one differencing round is sufficient to stabilize this variable with a p-value of 0.0109. The remaining variables, including SM, DIAB, and CVD, have p-values above 0.05, indicating that they are still non-stationary and require additional adjustment.

All variables become stationary at the second difference. For example, DIETR and SM become stationary with p-values less than 0.05, meanwhile CVD and DIAB have a p-values of 0.0137 and 0.000, respectively. This shows that two rounds of differencing are required for the majority of the variables in this dataset in order to prepare further analysis which is Multiple Linear Regression (MLR).

4.2 Risk Factors and Non-Communicable Diseases (NCDs)

Aligned with the first objective, this section examines the relationship between non-communicable diseases (NCDs) and their associated risk factors. The analysis includes descriptive statistics, correlation analysis and multiple linear regression (MLR). This section involves model 1 and model 2 of this research.

4.2.1 Descriptive Statistics

The descriptive statistics for Model 1 are displayed in table 3, which examines the prevalence of cardiovascular disease (CVD) and its linked risk factors which are dietary risk (DIETR), high body-mass index (HBMI), and smoking (SM).

Table 3
Descriptive statistics for Model 1

Variable	Cardiovascular disease (CVD)	Dietary risk (DIETR)	High body-mass index (HBMI)	Smoking (SM)
Mean	0.1866	0.0752	0.0499	0.0694
Median	0.1864	0.0743	0.0497	0.0697
Maximum	0.2011	0.081	0.0611	0.0710
Minimum	0.1754	0.0719	0.0407	0.061
Standard deviation	0.0089	0.0032	0.0060	0.0020
Skewness	0.3011	0.7229	0.2009	-3.4917
Kurtosis	1.7400	2.1103	1.8772	15.3868
Probability	0.4091	0.2670	0.5211	5.6468
Observations	22	22	22	22

For cardiovascular disease, the average prevalence is 0.1866, and the median value is 0.1864, suggesting that the data is consistent. The standard deviation is quite low at 0.0089, indicating minimal variability. While the kurtosis of 1.7400 suggests a flat distribution with fewer extreme cases, the skewness of 0.3011 reveals just a slight positive skew, indicating there are a few higher values. The prevalence of CVD fluctuates between a minimum of 0.1754 and a maximum of 0.2011, highlighting stability across the dataset.

For dietary risk, the mean level is 0.0752, with a median of 0.0743, showing limited variation as indicated by the standard deviation of 0.0032. The skewness of 0.7229 reflects a moderate tendency toward higher values, while the kurtosis of 2.1103 suggests that the data distribution is close to normal. Dietary risks range from 0.0719 to 0.0810, indicating consistency within the dataset.

The mean and median for the high body-mass index (HBMI) are 0.0499 and 0.0497, respectively, indicating consistency in the data. The low standard deviation of 0.0060 further supports the idea of stability in the data. The skewness of 0.2009 indicates a slight positive skew, while the kurtosis of 1.8772 suggests a relatively flat distribution without significant fluctuations.

The value of mean for smoking is 0.0694, while the standard deviation is relatively low which is 0.0020, indicating minimal variation. The skewness is highly negative at -3.4917, showing that most values are clustered at the lower end, while the high kurtosis of 15.3868 reveals a sharp concentration of values, likely due to a uniform pattern of smoking behaviour in the dataset.

Table 4
Descriptive statistics for Model 2

Variable	Cardiovascular disease (CVD)	Dietary risk (DIETR)	High body-mass index (HBMI)	Smoking (SM)
Mean	0.0315	0.0752	0.0499	0.0694
Median	0.0306	0.0743	0.0497	0.0696
Maximum	0.0385	0.0810	0.0611	0.0710
Minimum	0.0271	0.0719	0.0407	0.0610
Standard deviation	0.0031	0.0032	0.0060	0.0020
Skewness	0.6149	0.7229	0.2009	-3.4917
Kurtosis	2.3387	2.1103	1.8772	15.3868
Probability	0.4092	0.2669	0.5211	0.0000
Observations	22	22	22	22

Table 4 shows the descriptive statistics for Model 2, which examines the frequency of diabetes (DIAB) and its associated risk factors which are dietary risk (DIETR), high body-mass index (HBMI), and smoking (SM). For diabetes (DIAB), the average prevalence is 0.0315, with low variation reflected by a standard deviation of 0.0031. The median value is 0.0306, indicating a slight tendency toward lower values. The data is right skewed with a skewness of 0.6149, meaning higher values are more common. Additionally, the kurtosis of 2.3387 suggests that the data has a distribution closer to normal with no extreme values. The prevalence of diabetes ranges from a minimum of 0.0271 to a maximum of 0.0385, indicating a relatively narrow spread.

4.2.2 Correlation Analysis

Table 5 displays the correlation analysis for Model 1 which is the relationship of cardiovascular disease (CVD) and its linked risk factors.

Table 5
Correlation analysis for Model 1

	Cardiovascular disease (CVD)	Dietary risk (DIETR)	High body-mass index (HBMI)	Smoking (SM)
Cardiovascular disease (CVD)	1	0.97203	0.9411	-0.0648
Dietary risk (DIETR)	0.9703	1	0.9101	-0.0430
High body-mass index (HBMI)	0.9411	0.9101	1	-0.3672
Smoking (SM)	-0.0648	-0.0430	-0.3672	1

For cardiovascular disease (CVD) and dietary risk (DIETR), it indicates a strongest positive correlation between which valued 0.97203 followed by the second strongest correlation which is between cardiovascular disease (CVD) and high BMI (HBMI) which valued 0.9411. This explains that the severity of cardiovascular disease (CVD) is likely to increase in parallel with dietary risk (DIETR) or high BMI (HBMI). According to Xue *et al.*, [30], high-fat and high-sodium diets are well-established risk factors for cardiovascular diseases. Hence, these results are consistent with this medical literature. The strength of these correlations suggests a strong and direct link between these variables and the incidence of cardiovascular disease (CVD).

Besides, smoking had a weak negative correlation with cardiovascular disease (CVD) which valued -0.0648. According to this result, smoking does not appear to be strongly associated with cardiovascular disease in this particular dataset. The nature of smoking in this study group or other unaccounted-for confounding factors could be the cause of this poor connection according to Lushniak *et al*., [13]. Hence, one of the reasons for this could be the less intensity of smoking among the participant from this particular dataset that was obtained from Global Health Data Exchange (GHDX). Aside from the dataset, in one out of every four cardiovascular disease (CVD) deaths, smoking is the primary reason of the condition Centers for Disease Control and Prevention [5].

Table 6
Correlation analysis for Model 2

	Diabetes (DIAB)	Dietary risk (DIETR)	High body-mass index (HBMI)	Smoking (SM)
Diabetes (DIAB)	1	0.9147	0.9870	-0.3867
Dietary risk (DIETR)	0.9147	1	0.9101	-0.0430
High body-mass index (HBMI)	0.9870	0.9101	1	-0.3672
Smoking (SM)	-0.3867	-0.0430	-0.3672	1

For Model 2, the correlation analysis between diabetes and its risk factors is shown in Table 6. With a correlation coefficient of 0.9870, a strong positive correlation, the association between diabetes and high BMI is the most prominent in this model. This suggests that a person's risk of having diabetes rises significantly with an increase in BMI. According to European Society and Cardiology Press Office [9], Diabetes was 11 times more common in those with the highest BMI group than in those with the lowest BMI group. Therefore, maintaining a healthy weight and body-mass index is crucial in preventing diabetes.

Additionally, there is a significant positive relationship between dietary risk and diabetes, with a correlation of 0.9147. This suggests that the prevalence of diabetes in the study is strongly correlated with dietary risks. According to a study by Siddiqui *et al.,* [19], among those with pre-diabetes and undiagnosed diabetes, improved glycemic parameters were linked to a higher Healthy Eating Index (HEI) score, which indicates better food quality. Thus, unhealthy dietary habits, such as high consumption of sugar and fats, are key contributors to diabetes. Thus, this finding emphasizes the importance of promoting healthy eating to reduce the risk of diabetes.

However, smoking and diabetes have a negative correlation of -0.3867, indicating that smoking may not have as strong an association with diabetes compared to other factors like high BMI and dietary risk. However, smokers are 30–40% more likely to develop diabetes compared to non-smokers according to Center for Disease Control and Prevention [4]. This finding highlights the impact of smoking, which not only raises the risk of cardiovascular and respiratory conditions but also contributes to diabetes development.

4.2.3 Variance Inflation Factor (VIF)

Since the correlation values of the independent variables for both models 1 and 2 are highly correlated, regression coefficient estimations may be unstable and incorrect. The variance inflation factor (VIF) test is therefore helpful in determining how severe multicollinearity is. Table 7 and 8 shows the variance inflation factor (VIF) value for independent variables for model 1 and 2 with and without one of the independent variables which is DIETR respectively.

Table 7
Variance Inflation Factor (VIF) analysis for model 1 and model 2

Variable	Variance Inflation Factor (VIF) Value
Dietary Risk (DIETR)	13.5451
High Body-Mass Index (HBMI)	15.6215
Smoking (SM)	2.6656

Table 8
Variance Inflation Factor (VIF) analysis for model 1 and model 2 without DIETR

Variable	Variance Inflation Factor (VIF) Value
High Body-Mass Index (HBMI)	1.1557
Smoking (SM)	1.1557

For dietary risk (DIETR), the value of variance inflation factor (VIF) is 13.5451 and it indicates a serious multicollinearity problem since the VIF value exceeds 10. This explains a substantial amount of variance between DIETR and other variables such as HBMI and SM in these models. For high body-mass index (HBMI), the value of VIF is 15.6215 indicates a severe multicollinearity. This higher VIF value indicates a stronger relationship between HBMI and other independent variables. A higher multicollinearity leads to an unreliable regression coefficient.

In order to solve this matter, DIETR was removed from the VIF analysis and a new calculation for the VIF values were made and represents in Table 8. The new VIF values for both of the independent variables decrease to a value below 5 and it shows moderate multicollinearity, which is tolerable and not severe. Both high body-mass index (HBMI) and smoking (SM) have a VIF value of 1.1557. Thus, the regression coefficients for HBMI and SM are now more dependable and more reliable and accurately reflect to the dependent variable for each of model 1 and model 2, which is cardiovascular disease (CVD) and diabetes (DIAB) respectively.

4.2.4 Multiple Linear Regression (MLR)

Table 9 shows the regression analysis for model 1 that examines how dietary risk (DIETR), high body-mass index (HBMI) and smoking (SM) contribute to the prevalence of cardiovascular disease (CVD) meanwhile table 10 shows the analysis after removing one of the independent variables which is DIETR.

Table 9
Regression analysis for Model 1

Variable	Coefficient	p-value
Dietary Risk (DIETR)	0.7264	0.0381
High Body-Mass Index (HBMI)	1.1660	0.0000
Smoking (SM)	1.0407	0.0003
Constant	0.0015	0.9043

Table 10
Regression analysis for Model 1 without DIETR

Variable	Coefficient	p-value
High Body-Mass Index (HBMI)	1.5568	0.0000
Smoking (SM)	1.4195	0.0000
Constant	0.0098	0.4493

For the analysis that includes dietary risk (DIETR) in table 11, all of the independent variables show a significant value associated with the dependent variable, diabetes (DIAB) and high body-mass index (HBMI) shows the strongest association with coefficient of 0.3186 and p-value of 0.0110. Moreover, dietary risk (DIETR) also has a significant association with CVD with coefficient of 0.3186 and p-value of 0.0110. According to Beigrezaei *et al*., [2], some dietary patterns, especially those high in fruits, were linked to a lower risk of type 2 diabetes, indicating that diet has a big impact on diabetes risk. Conversely, smoking (SM) has a significant negative relationship with DIAB with coefficient of -0.2125 and 0.0149.

In addition, after excluding DIETR in addressing multicollinearity issue, HBMI remains as a significant variable and still positively associated with DIAB. The new value of coefficient and p-value for HBMI are 0.5111 and 0.0000, respectively. However, smoking becomes insignificant with a new value of coefficient and p-value for SM are -0.0442 and 04757, respectively. Hence, HBMI is resulted as the strongest positive association with diabetes (DIAB). According to Bai *et al*., [1], similar findings were seen for both genders, where a greater BMI was linked to a higher risk of diabetes.

In contrary, for smoking (SM), shows a negative association with CVD. This finding might seem impossible, it might represent underlying complications or particular patterns of behavior in the population under study. However, few of the past studies proved the contrary of this result. according to Lushniak *et al*., [13], type 2 diabetes is between 30% and 40% more common in smokers than in non-smokers. Besides, smokers, even those who are not overweight, are more likely to have belly fat, which increases the risk of type 2 diabetes according to Centers for Disease Control and Prevention [4]. Hence, these studies emphasize that smoking rises the risk of developing diabetes. To conclude for model 2, H4 and H5 were supported meanwhile H6 were unsupported. Thus, to answer the first objective, high BMI have the most effects on each of the non-communicable diseases (NCDs) under study, namely diabetes and cardiovascular diseases.

4.3 Non-Communicable Diseases (NCDs) and Healthcare Expenditure

Aligned with the second objective, this section explores the relationship between the non-communicable diseases and how it affects the healthcare expenditure in Malaysia. This section involves model 3 of this research.

4.3.1 Descriptive statistics

Table 13 displays the descriptive statistics for Model 3, which investigates the relationship between healthcare expenditure (HEX), cardiovascular disease (CVD), and diabetes (DIAB).

Table 13
Descriptive Statistics for Model 3

Variable	Dietary risk (DIETR)	High body-mass index (HBMI)	Smoking (SM)
Mean	290.0467	0.1866	0.0315
Median	314.1126	0.1864	0.0306
Maximum	487.0088	0.2010	0.0385
Minimum	111.9489	0.1754	0.0271
Standard deviation	120.0433	0.0089	0.0031
Skewness	-0.1647	0.3011	0.6149
Kurtosis	1.6283	1.7400	2.3387
Probability	0.4017	0.4091	0.4092
Observations	22	22	22

The mean of healthcare expenditure (HEX) is 290.0467 million US dollars, with a standard deviation of 120.0433 million, indicating significant variability in spending. The skewness of -0.1647 suggests that the data is fairly symmetrical, while the kurtosis of 1.6828 reveals a flatter distribution compared to a normal curve. Healthcare spending ranges from a minimum of 111.9489 million US dollars to a maximum of 487.0088 million US dollars, reflecting substantial differences in the data.

The mean prevalence of cardiovascular disease is 0.1855, with a median of 0.1864, indicating low variability. The skewness of 0.3011 shows a slight positive skew, while the kurtosis of 1.7400 suggests a distribution with fewer extreme values. The prevalence of CVD fluctuates between a minimum of 0.1754 and a maximum of 0.2011, highlighting stability within the data.

The mean prevalence of diabetes is 0.0315, with a standard deviation of 0.0031, showing moderate variability. The skewness of 0.6149 indicates a slight positive skew, suggesting that the distribution is concentrated on higher values. The kurtosis of 2.3387 shows a mesokurtic distribution, meaning the data is closer to a normal distribution. The prevalence of diabetes ranges from 0.0271 to 0.0385, reflecting a narrower spread compared to healthcare expenditure.

4.3.2 Correlation analysis

Table 14 represents the correlation analysis for healthcare expenditure (HEX) and the two NCDs, which are cardiovascular disease (CVD) and diabetes (DIAB).

Table 14
Correlation analysis for Model 3

	Healthcare Expenditure (HEX)	Diabetes (DIAB)	Cardiovascular Disease (CVD)
Healthcare Expenditure (HEX)	1	0.8191	0.8812
Diabetes (DIAB)	0.8191	1	0.9100
Cardiovascular Disease (CVD)	0.8812	0.9100	1

With a correlation of 0.8812, the relationship between healthcare expenditures and cardiovascular disease is highly significant and positive, indicating that increased healthcare spending is associated with the higher prevalence and severity of cardiovascular disease. In Malaysia, cardiovascular illnesses cost more than RM 3.93 billion annually, which makes up 40.73% of the total expenses for certain non-communicable diseases according to World Health Organization [27].

Besides, there is a strong positive association between diabetes and healthcare expenditures, with a correlation value of 0.8191. This suggests that as the prevalence of diabetes increases, so does healthcare spending, reflecting the significant financial burden of managing diabetes in Malaysia. According to Ganasegeran *et al*., [10], diabetes costs Malaysia an estimated USD 600 million annually, demonstrating its substantial economic impact. Additionally, the World Health Organization [16], stated that the yearly medical expenses for cancer, diabetes, and cardiovascular illnesses in Malaysia surpass RM 9.65 billion, further emphasizing the critical financial impact of these non-communicable diseases.

4.3.3 Variance Inflation Factor (VIF)

Table 15 displays the value of variance inflation factor (VIF) for independent variables for model 3.

Table 15
Variance Inflation Factor (VIF) analysis for Model 3

Variable	Variance Inflation Factor (VIF) Value
Cardiovascular Disease (CVD)	5.8160
Diabetes (DIAB)	5.8160

For cardiovascular disease (CVD) and diabetes (DIAB), the VIF value is 5.8160 and it indicates a moderate multicollinearity. This indicates that there is some correlation between CVD and DIAB, which will increase the regression coefficient's variability.

4.3.4 Multiple Linear Regression (MLR)

Table 16 shows the regression analysis for model 3 that examines the impact of cardiovascular disease (CVD) and diabetes (DIAB) on healthcare expenditure (HEX).

Table 16
Regression analysis for Model 3

Variable	Coefficient	p-value
Cardiovascular Disease (CVD)	1360.502	0.7042
Diabetes (DIAB)	30201.79	0.0069
Constant	-913.6081	0.0032

For cardiovascular disease (CVD), the coefficient is 1360.502 with a p-value of 0.7042, indicating no significant relationship between CVD and HEX. Although this result does not align with expectations, previous studies have consistently demonstrated the significant financial burden of CVD. According to the Ministry of Health Malaysia [16], the annual cost of treating cardiovascular disease in Malaysia is approximately MYR 3.9 billion, which accounts for over 40% of all direct healthcare expenses related to non-communicable diseases. These findings imply that even while the regression model may not account for CVD's influence, its wider impact on the economy is still significant.

Besides, for diabetes (DIAB), the coefficient is 30201.79 with a p-value of 0.0069, representing a strong positive relationship with healthcare expenditure. This suggests that higher diabetes prevalence significantly increases healthcare costs. Previous studies support this finding, highlighting the considerable economic burden of diabetes in Malaysia. For instance, Ganasegeran *et al.*, [10], estimated the total yearly cost of diabetes in Malaysia to be approximately USD 600 million. Similarly, Sun *et al.*, [21], reported direct healthcare costs related to diabetes amounting to RM 4.4 billion annually. To conclude, H7 and H8 were supported. Hence, to answer the second objective, diabetes contributes for the most portion of Malaysia’s healthcare expenditures. To solve this problem, Malaysia’s government need to improve the efficient management in reducing the financial strain on healthcare system and also the need of Ministry of Health in more effective prevention techniques.

5. Conclusion

In summary, this study effectively addressed the research question by looking at how risk factors affect non-communicable disease and how non-communicable diseases (NCDs) affect Malaysia's health care expenditures between 2000 and 2021. Risk factors like smoking, high body-mass index (BMI), and dietary risk are identified in the analysis as major causes of NCDs like diabetes and cardiovascular disease. Among the risk factors, high BMI were highly affected both of the diseases which are cardiovascular disease (CVD) and diabetes (DIAB) and diabetes is the highest contributor to healthcare expenditure in Malaysia. To reduce the financial strain on the healthcare system, the study emphasizes the necessity of efficient management and prevention techniques.

This study has limitations regardless of these findings. The use of secondary data, which might not fully capture all factors influencing healthcare expenditures, is one of the significant limitations. Furthermore, the indirect costs such as lost productivity or long-term care were not examined in this study, even though they would offer a more complete picture of the financial burden of NCDs. Another limitation is the exclusion of other socio-economic and environmental aspects that may contribute to the rising occurrence of NCDs in Malaysia.

Recommendation from this research is for the future studies to observe both the direct and indirect costs of non-communicable diseases which include long-term care, productivity loss, and the wider socioeconomic effects, to address these issues. It is also recommended that policymakers focus on strengthening prevention programs, particularly for diabetes and cardiovascular disease, to reduce the long-term financial burden on the healthcare system.

References

[1] Bai, Kaizhi, Xuejiao Chen, Rui Song, Wenlong Shi, and Songhe Shi. "Association of body mass index and waist circumference with type 2 diabetes mellitus in older adults: a cross-sectional study." *BMC geriatrics* 22, no. 1 (2022): 489. https://doi.org/10.1186/s12877-022-03145-w

[2] Beigrezaei, Sara, Sara Jambarsang, Sayyed Saeid Khayyatzadeh, Masoud Mirzaei, Amir Houshang Mehrparvar, and Amin Salehi-Abargouei. "The association between dietary patterns derived by three statistical methods and type 2 diabetes risk: YaHS-TAMYZ and Shahedieh cohort studies." *Scientific Reports* 13, no. 1 (2023): 410. https://doi.org/10.1038/s41598-023-27645-w

[3] Catalyst, N. E. J. M. (2018). What is risk management in healthcare?. *NEJM Catalyst, 4(2)*.

[4] Centers for Disease Control and Prevention. (2022). *Smoking and diabetes*. CDC. https://www.cdc.gov/tobacco/campaign/tips/diseases/diabetes.html

[5] Centers for Disease Control and Prevention. (2024). *Health effects of cigarettes: cardiovascular disease.* CDC. https://www.cdc.gov/tobacco/about/cigarettes-and-cardiovascular-disease.html

[6] Centers for Disease Control and Prevention. (2024). *Heart Disease Risk Factors*. CDC. https://www.cdc.gov/heart-disease/risk-factors/index.html

[7] Chandran, Arunah, Shurendar Selva Kumar, Noran Naqiah Hairi, Wah Yun Low, and Feisul Idzwan Mustapha. "Non-communicable disease surveillance in Malaysia: An overview of existing systems and priorities going forward." *Frontiers in Public Health* 9 (2021): 698741.

[8] Diana, R., M. Safrin, and A. M. Rosliza. "Management of Non-Communicable disease prevention at district level in Malaysia." *International Journal of Public Health and Clinical Sciences* 6, no. 4 (2019): 53-67.

[9] European Society of Cardiology Press Office. (2020). *Body mass index is a more powerful risk factor for diabetes than genetics.* ESC. https://www.escardio.org/The-ESC/Press-Office/Press-releases/Body-mass-index-is-a-more-powerful-risk-factor-for-diabetes-than-genetics

[10] Ganasegeran, Kurubaran, Chee Peng Hor, Mohd Fadzly Amar Jamil, Hong Chuan Loh, Juliana Mohd Noor, Norshahida Abdul Hamid, Purnima Devi Suppiah, Mohd Rizal Abdul Manaf, Alan Swee Hock Ch'ng, and Irene Looi. "A systematic review of the economic burden of type 2 diabetes in Malaysia." *International journal of environmental research and public health* 17, no. 16 (2020): 5723. https://doi.org/10.3390/ijerph17165723

[11] International Federation of Red Cross and Red Crescent Societies. (n.d.). *Non-communicable disease*. IFRC. https://www.ifrc.org/our-work/health-and-care/community-health/non-communicable-diseases

[12] Lim, Ooi Wei. "Risk factors of non-communicable diseases in Malaysia/Lim Ooi Wei." PhD diss., Universiti Malaya, 2019.
[13] Lushniak, Boris D., Jonathan M. Samet, Terry F. Pechacek, Leslie A. Norman, and Peter A. Taylor. "The health consequences of smoking—50 years of progress: A report of the surgeon general." (2014).
[14] Ministry of Health Malaysia. (2017). *Health cost of non-communicable diseases (NCDs) in Malaysia*. https://www.moh.gov.my/moh/resources/Penerbitan/Rujukan/NCD/NCD_Laporan/HEALTH-COST_of_NCDs-7a-WEB.pdf
[15] Ministry of Health Malaysia. (2020). *The Impact of Noncommunicable Diseases and Their Risk Factors on Malaysia's Gross Domestic Product.* https://www.moh.gov.my/index.php/pages/view/1904
[16] Ministry of Health Malaysia. (2022). *The Direct Health-Care Cost of Noncommunicable Diseases in Malaysia*. https://www.moh.gov.my/moh/resources/Penerbitan/Rujukan/NCD/NCD_Laporan/HEALTH-COST_of_NCDs-7a-WEB.pdf
[17] NHS UK. (2024). *Cardiovascular disease*. nhs.uk. https://www.nhs.uk/conditions/cardiovascular-disease/
[19] Siddiqui, Sania, Hadzliana Zainal, Sabariah Noor Harun, and Siti Maisharah Sheikh Ghadzi. "Diet quality and its association with glycemic parameters in different diabetes progression stages. A cross-sectional questionnaire study at a primary care clinic." *Clinical Nutrition ESPEN* 39 (2020): 165-172. https://doi.org/10.1016/j.clnesp.2020.06.022
[20] Stanford University. (n.d.). *Definition of risk*. Office of the Chief Risk Officer. https://ocro.stanford.edu/enterprise-risk-management-erm/key-definitions/definition-risk
[21] Sun, H., P. Saeedi, S. Karuranga, M. Pinkepank, K. Ogurtsova, B. B. Duncan, C. Stein et al. *IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabe-tes Res Clin Pract 183: 109119*. 2021.
[22] The Investopedia Team. (2024). *Variance Inflation Factor (VIF*). Investopedia. https://www.investopedia.com/terms/v/variance-inflation-factor
[25] World Cancer Research Fund International. (2014). *The Link between Food, Nutrition, Diet and Non-Communicable Diseases*. https://www.wcrf.org/wp-content/uploads/2024/12/WCRF-NCD-A4-WEB.pdf
[26] World Health Organization. (2021). *Cardiovascular diseases (CVDs).* https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
[27] World Health Organization. (2022). *The annual health-care cost of cardiovascular diseases, diabetes and cancer in Malaysia exceeds RM 9.65 billion.* https://www.who.int/malaysia/news/detail/09-08-2022-the-annual-health-care-cost-of-cardiovascular-diseases--diabetes-and-cancer-in-malaysia-exceeds-rm-9.65-billion
[28] World Health Organization. (2024). *Diabetes*. https://www.who.int/news-room/fact-sheets/detail/diabetes
[29] World Health Organization. (n.d.). *Noncommunicable diseases: Risk factors and conditions*. https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/ncd-risk-factors
[30] Xue, Xiao-Dan, Wei Li, Mei-Qiu Xie, De-Zheng Wang, Dan-Dan Li, Peng Xin, Wen-Long Zheng, and Guo-Hong Jiang. "High sodium diet intake and cardiovascular diseases: An attributable death study in Tianjin, China." *The Journal of Clinical Hypertension* 25, no. 1 (2023): 30-37. https://doi.org/10.1111/jch.14599