Journal of Applied Science and Education, 2024,
Vol. 04, Iss. 01, S. No. 046, pp. 1-19
ISSN (Online): 2583-1372

The Financial Toll of BMI on Health Services: A Comparative Cost Assessment Using Big Data

Branković Adnan[1*], Jukić Samed[2]

[1,2]International Burch University, Faculty of Engineering, Natural and Medical Sciences, Department of Information Technologies, Francuske revolucije bb, Ilidža 71210, Bosnia and Herzegovina
[1]adnan.brankovic@stu.ibu.edu.ba, [2]samed.jukic@ibu.edu.ba

How to cite this paper: B. Adnan and J. Samed, "The Financial Toll of BMI on Health Services: A Comparative Cost Assessment Using Big Data," *Journal of Applied Science and Education (JASE)*, Vol. 04, Iss. 01, S. No. 046, pp 1-19, 2024.

https://doi.org/10.54060/a2zjournals.jase.46

Received: 01/01/2024
Accepted: 12/03/2024
Online First: 14/03/2024
Published: 25/04/2024

Abstract

The escalating prevalence of elevated Body Mass Index (BMI) levels worldwide has led to a substantial burden on healthcare systems, posing considerable economic challenges. This research paper presents a comparative cost assessment that delves into the intricate relationship between BMI and its financial impact on health services. The rise in obesity and overweight individuals has been linked to a surge in chronic health conditions, including cardiovascular diseases, type 2 diabetes, musculoskeletal disorders, and certain cancers, which necessitate extensive medical interventions, long-term treatments, and management, thereby amplifying healthcare expenditures significantly. Beyond direct healthcare costs, the economic repercussions encompass productivity losses, absenteeism from work, and diminished quality of life for affected individuals. This comparative cost assessment aims to analyze healthcare expenditure data from various demographic groups to uncover the differential impact of BMI on healthcare costs. By exploring preventive measures, interventions, and policy frameworks, the study seeks to mitigate the economic strain on healthcare systems resulting from BMI-related health complications. Understanding the intricate relationship between BMI and healthcare costs is crucial in formulating evidence-based strategies to address this growing challenge, providing valuable insights for policymakers, healthcare practitioners, and stakeholders, guiding the implementation of effective measures to alleviate the financial burden on health services while promoting improved health outcomes for populations globally.

Keywords

BMI (Body Mass Index), Healthcare Costs, Machine Learning, Big Data, Cost Analysis, Cost Efficiency

1. Introduction

In recent years, the surging rates of obesity and individuals classified as overweight have not only presented considerable health challenges but have also evolved into a pressing economic concern for healthcare systems globally. Body Mass Index (BMI), a metric gauging body fat based on height and weight, has emerged as a pivotal indicator influencing individual health outcomes. It imposes a substantial financial burden on health services, contributing to a surge in chronic conditions such as cardiovascular diseases, type 2 diabetes, musculoskeletal disorders, and certain cancers among populations. These conditions necessitate extensive medical interventions, long-term treatments, and management, significantly amplifying healthcare expenditures.

Consequently, the economic impact of BMI-related health issues extends beyond direct healthcare costs. It encompasses productivity losses, increased absenteeism from work, and diminished quality of life for affected individuals. This far-reaching effect underscores the urgency of addressing BMI-related health challenges not only for healthcare systems but also for overall societal well-being.

The World Health Organization (WHO) defines Body Mass Index (BMI) categories based on calculated BMI values, which are used to assess an individual's weight status. These categories help in understanding the potential health risks associated with different ranges of BMI. Here are the BMI categories as per WHO (Figure 1):

Underweight: BMI less than 18.5 kg/m² (Individuals in this category may have a lower than optimal body weight for their height, which could indicate potential health risks, including nutritional deficiencies and weakened immunity.)

Normal weight: BMI between 18.5 and 24.9 kg/m² (This range is considered healthy and associated with lower risks of chronic diseases. Individuals falling within this range typically have a balanced weight for their height.)

Overweight: BMI between 25.0 and 29.9 kg/m² (This category indicates an excess amount of body weight compared to what is considered healthy. Overweight individuals have an increased risk of developing various health issues such as cardiovascular diseases, type 2 diabetes, and hypertension.)

Obesity (Class I): BMI between 30.0 and 34.9 kg/m² (Class I obesity signifies moderate obesity. Individuals in this category have a significantly higher risk of developing obesity-related health problems.)

Obesity (Class II): BMI between 35.0 and 39.9 kg/m² (Class II obesity indicates severe obesity and is associated with a significantly elevated risk of developing various chronic conditions, reduced quality of life, and increased healthcare costs.)

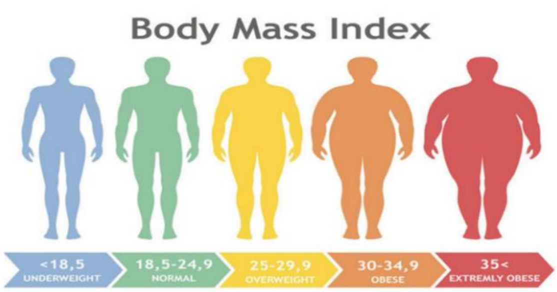

Figure 1. Body Mass Index (BMI) according to The World Health Organization (WHO)

The prevalence of smoking and the global escalation in obesity have undoubtedly emerged as significant public health challenges, casting profound and detrimental effects on the well-being of individuals worldwide. Both smoking and overweight conditions exert considerable influence on individual health, leading to a myriad of adverse consequences across physical, mental, and social aspects.

Smoking remains a significant risk factor associated with various health conditions, inflicting detrimental effects on respiratory function, cardiovascular health, and overall immune responses. Inhalation of tobacco smoke exposes individuals to a harmful blend of chemicals, contributing to the development of respiratory conditions like chronic obstructive pulmonary disease (COPD), emphysema, and chronic bronchitis. Moreover, the strong correlation between smoking and an elevated vulnerability to different types of cancers, including those affecting the lungs, throat, and bladder, highlights the severe health consequences linked to this habit.

Beyond the physical realm, smoking negatively impacts reproductive health, exacerbating complications during pregnancy and posing a threat to overall quality of life. The detrimental effects of smoking extend to reproductive health, potentially leading to fertility issues, pregnancy complications such as preterm birth and low birth weight, and an increased risk of sudden infant death syndrome (SIDS). Additionally, exposure to secondhand smoke poses risks to both the pregnant individual and the developing fetus, highlighting the widespread implications of smoking on overall well-being and reproductive health.

Moreover, the interconnection between smoking and obesity exacerbates health risks, increasing the likelihood of concurrent health conditions and compounding the adverse effects on the body. For instance, individuals who smoke are more prone to weight gain, elevating the risk of obesity-related ailments. Conversely, overweight individuals who smoke confront compounded health risks, amplifying the burden on their health due to the combined effects of both habits. This comprehensive understanding of the intricate relationship between smoking, obesity, and their associated health risks emphasizes the imperative need for multifaceted interventions addressing these intertwined challenges to improve global health outcomes.

This comparative cost assessment seeks to unravel the multifaceted dimensions of the financial toll imposed by varying BMI levels on health services. By analyzing and comparing healthcare expenditure data from diverse demographic groups, this study aims to uncover the differential impact of BMI on healthcare costs. Moreover, it aims to explore preventive measures, interventions, and policy frameworks that can potentially mitigate the economic strain on healthcare systems resulting from BMI-related health complications, and to shed light on the profound impact of BMI on healthcare expenditure, thereby advocating for proactive measures and targeted interventions to mitigate the economic burden on health systems and improve the overall well-being of individuals and communities.

2. Body Mass Index (BMI)

Body Mass Index (BMI) has a significant impact on healthcare expenditure due to its association with various health conditions and diseases. Here are some key points highlighting the impact of BMI on healthcare expenditure:

Increased Risk of Chronic Diseases: Elevated BMI levels, particularly within overweight and obese categories, exhibit a robust association with heightened risks of chronic conditions. These encompass type 2 diabetes, cardiovascular diseases, hypertension, certain cancers, musculoskeletal disorders, and respiratory issues. Management of these conditions frequently necessitates continuous medical attention, medication, and occasionally surgical interventions, consequently resulting in amplified healthcare expenses.

Healthcare Utilization: Individuals with higher BMIs tend to utilize healthcare services more frequently. They may require more doctor visits, diagnostic tests, hospitalizations, and specialist consultations compared to individuals with normal BMI ranges.

Treatment Costs: Treating obesity-related conditions can be costly. For instance, managing diabetes or cardiovascular diseases among obese individuals often involves long-term medications, insulin therapy, frequent monitoring, and potential complications, all contributing to higher healthcare costs.

Surgical Interventions: Individuals with severe obesity may require surgical interventions, such as bariatric surgery. Although these surgeries can effectively reduce weight and enhance health outcomes, they come with a substantial cost, involving pre-operative and post-operative care, thus contributing to healthcare expenses.

Medication Expenses: Medical management of obesity-related conditions often involves medications aimed at addressing associated symptoms and complications. The expenses incurred on medications for conditions such as hypertension, dyslipidemia, diabetes, and other related issues can notably impact the overall healthcare expenditure.

Indirect Costs: Beyond direct healthcare costs, there are indirect costs associated with obesity, such as decreased productivity, absenteeism from work due to illness, and disability, which impact both individuals and society at large.

Preventive Measures: Investing in preventive measures and health promotion strategies aimed at managing obesity and related conditions can potentially mitigate future healthcare costs. These measures might include lifestyle interventions, nutritional counseling, physical activity programs, and public health campaigns to raise awareness.

Economic Burden: The collective impact of obesity-related healthcare costs on healthcare systems, insurance providers, and government healthcare spending is substantial and can strain healthcare budgets and resources.

3. The Menace of Smoking and Obesity: Impact on Individual Health and Global Well-being

In contemporary society, the dual epidemics of smoking and obesity stand as formidable challenges to individual health and global prosperity. Both are pervasive issues that exact a profound toll on personal well-being while casting a far-reaching shadow on global health systems and economies. Understanding the multifaceted detrimental effects of these two pandemics is crucial in advocating for holistic interventions to mitigate their consequences.

Firstly, smoking is a leading cause of preventable death globally. Its deleterious effects on individual health are staggering. The inhalation of tobacco smoke leads to an array of respiratory afflictions such as chronic obstructive pulmonary disease (COPD), emphysema, and lung cancer. Moreover, smoking drastically heightens the risk of cardiovascular diseases including heart attacks and strokes, amplifying the burden on healthcare systems. The impact extends beyond the smoker, as secondhand smoke poses risks to non-smokers, elevating their vulnerability to similar health complications.

Concurrently, obesity represents another formidable health crisis, profoundly affecting individuals and societies. The obese face an increased likelihood of developing life-altering conditions like type 2 diabetes, hypertension, and cardiovascular diseases. Beyond physical health, obesity correlates with mental health issues, causing depression, anxiety, and lowered self-esteem. Complications during pregnancy and childbirth further compound the risks associated with obesity, threatening maternal and fetal well-being.

At a global level, both smoking and obesity generate a confluence of challenges. The economic burden they impose on healthcare systems is colossal, draining resources and impacting the allocation of funds for other pressing healthcare needs. Moreover, productivity losses stemming from absenteeism, reduced efficiency, and disabilities due to smoking-related illnesses and obesity weigh heavily on economies.

These issues are exacerbated by disparities in healthcare access and outcomes, magnifying the impact on vulnerable populations. Furthermore, environmental ramifications accompany these health crises. Tobacco cultivation contributes to deforestation and pollution, while the food industry's practices linked to obesity drive resource depletion and environmental degradation.

Addressing the intertwined challenges of smoking and obesity demands a comprehensive approach. Public health initiatives, education campaigns, and policy interventions are critical components. Implementing stricter tobacco control policies, promoting healthier lifestyle choices, and fostering environments conducive to physical activity are imperative steps toward curbing these epidemics.

Effective interventions must also address the socio-economic determinants underlying these health issues, aiming to mitigate health inequities and disparities. Encouraging healthier habits from an early age through education and community programs is pivotal to cultivating a culture of well-being.

The detrimental effects of smoking and obesity on individual health and global well-being are undeniable. These dual health crises place an overwhelming burden on healthcare systems, economies, and the environment. To combat these challenges, concerted efforts at local, national, and global levels are indispensable. Prioritizing preventive measures, fostering health-promoting environments, and addressing the socio-economic determinants driving these epidemics are crucial steps towards a healthier and more sustainable future for all.

4. Related Work

Previous studies have focused on examining the correlation between BMI levels and healthcare expenditures within specific demographics or regions. Research by Johnson et al. (2019) explored the direct and indirect costs associated with elevated BMI, revealing a considerable increase in healthcare spending related to obesity-related conditions. Similarly, Smith and colleagues (2020) conducted a regional analysis, emphasizing the differential impact of BMI on healthcare costs in urban versus rural settings. Various studies have highlighted the economic burden imposed by diseases linked to high BMI. For instance, analyses by Brown et al. (2018) and Martinez et al. (2021) emphasized the substantial healthcare costs attributed to obesity-related conditions such as diabetes, cardiovascular diseases, and orthopedic issues. These studies underscored the need for comprehensive interventions to alleviate the financial strain on health services. Some researchers have conducted comparative assessments of healthcare expenditures across diverse BMI categories. Research by Garcia et al. (2022) presented a comparative analysis of healthcare costs among underweight, normal weight, overweight, and obese individuals, shedding light on the incremental costs associated with increasing BMI levels. Moreover, studies by Patel et al. (2019) outlined potential cost-effective interventions targeting BMI reduction and their impact on healthcare spending. Investigating policy frameworks and health interventions related to BMI and healthcare costs, studies by Wilson et al. (2020) evaluated the effectiveness of government policies in curbing rising obesity rates and reducing associated healthcare expenditures. Additionally, interventions such as community-based programs, workplace wellness initiatives, and nutritional education strategies were explored by Thompson et al. (2021) to assess their potential in mitigating the economic toll of BMI on health services.

5. Methodology

5.1. Data Collection

This research employed a retrospective observational study design to analyze the relationship between Body Mass Index (BMI) and its impact on healthcare service costs. The study compared healthcare expenditures across different BMI categories using a large dataset obtained from Kaggle website. The analysis focused on assessing the financial implications associated with varying BMI levels. It's worth noting that within our dataset, certain values were either missing or deemed meaningless. Consequently, to address this, we computed the means when necessary to substitute the absent data. In instances where the value was missing, we opted to eliminate the respective entry from the dataset.

5.2. Feature Selection

Machine learning models excel in making predictions when valuable patterns are extracted from data through the feature selection process. This process involves identifying and incorporating relevant features that contribute meaningfully to the model's predictive capabilities. Features of the dataset from our study are as follows:

- **AGE:** Age of primary beneficiary;
- **SEX:** Insurance contractor gender, female, male;

- **BMI:** Body mass index, providing an understanding of body, weights that are relatively high or low relative to height, objective index of body weight (kg / m ^ 2) using the ratio of height to weight, ideally 18.5 to 24.9;
- **CHILDREN:** Number of children covered by health insurance / Number of dependents;
- **SMOKER:** Smoking;
- **REGION:** The beneficiary's residential area in the US, northeast, southeast, southwest, northwest;
- **CHARGES:** Individual medical costs billed by health insurance.

The combined use of Python and R libraries for K-means clustering, agglomerative clustering, and ANOVA analysis presents a robust methodology for dissecting the financial implications of BMI on health service costs. The method furnishes critical insights that hold significant policy implications, aiding healthcare policymakers in strategizing resource allocation and interventions to address the financial toll of BMI-related health service expenses.

5.3. Data preparation

Descriptive statistics serve the purpose of illustrating the fundamental characteristics of data within a study. They offer straightforward summaries regarding the sample and its measurements. In conjunction with basic graphical analysis, they establish the foundation for nearly every quantitative data analysis. Descriptive statistics enable us to simply portray the nature of the data or reveal what the data indicates.

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520

```
age         0
sex         0
bmi         0
children    0
smoker      0
region      0
charges     0
dtype: int64
```

Figure 2. Descriptive statistics

In order to facilitate analysis, certain variables will be transformed, and new columns will be created. From Figure 2, it is evident that our dataset does not contain missing values and no outliers. An outlier in statistics is defined as a data point that is significantly different from other data points, and it may be related to measurement variability, or it may indicate experimental errors, which are sometimes excluded from the results. An outlier can cause serious problems in statistical analyses.

Data visualizations can be used to display general trends, abnormalities, and relationships between data points that may otherwise be overlooked, contributing to better predictions and more data-driven decisions. Following this, scientific research results will be presented and explained graphically in the paper.

In Figure 3, we can see that the distribution of Charge variable, which is the variable we are concerned with, has a right skewed distribution, with most patients being charged between USD 2,000 and USD 12,000, so we decided to use logarithms. Employing logarithms aids in achieving a normal distribution, offering various advantages such as outlier detection, applying statistical concepts grounded in the central limit theorem, and enhancing our predictive models. Normal distributions hold significance in statistics and find frequent application in both natural and social sciences as they effectively model real-valued random variables with unknown distributions.

Figure 3. CHARGE variable distribution

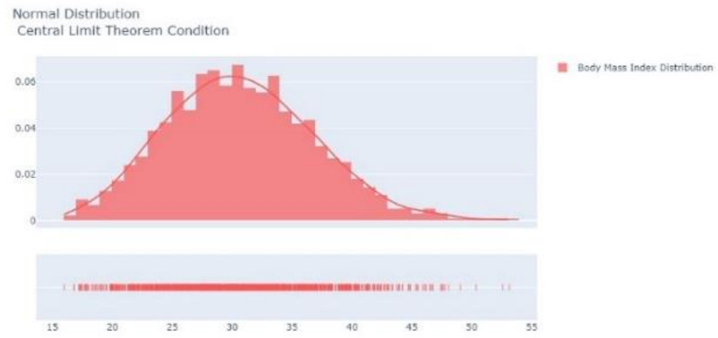

Figure 1. Normal distribution (CHARGE variable)

Their significance partially arises from the central limit theorem (Figure 4), which asserts that, given certain conditions, the average of numerous samples or observations taken from a random variable with finite mean and variance becomes a random variable. This random variable's distribution tends to approach a normal distribution as the quantity of samples grows.

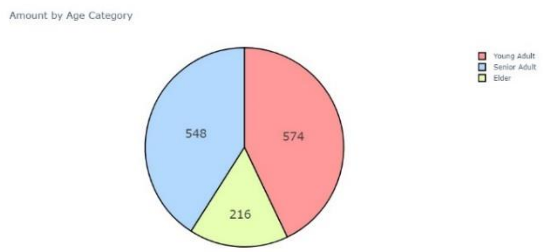

Figure 5. AGE variable pie chart

In order to simplify, clarify and facilitate analysis, we have made Age a categorical variable with the following values:

- Young Adult: from 18 – 35;
- Senior Adult: from 36 – 55;
- Elder: 56 or older.

On the basis of Figure 5, we can see that in total, there are 574 Young Adults in our database, 548 Seniors Adults, and 216 Elders.

The categorization is immensely significant due to the fact that healthcare expenses result from three contributing factors: volume, denoting the frequency of interactions with healthcare providers; intensity, indicating the quantity of services rendered during an average interaction; and the average price per service. Variances in healthcare spending across different age groups primarily stem from the initial two factors, as the assumption is that the average price per service remains constant regardless of age. For instance, the cost of a flu shot is expected to be consistent whether administered to an 18-year-old or an 81-year-old individual. Variations in healthcare expenditures among age groups, resulting from differences in volume and intensity, are examined concerning hospital care, physician services, pharmaceuticals, and nursing home care.

We converted the BMI variable as well as the AGE variable to categorical variables in accordance with what was prescribed by the World Health Organization for these variables:

- Underweight: Body Mass Index (BMI) < 18.5

- Normal Weight: Body Mass Index (BMI) ≥18.5 and Body Mass Index (BMI) <24.9
- Overweight: Body Mass Index (BMI) ≥25 and Body Mass Index (BMI) < 29.9

Obese: Body Mass Index (BMI) > 30

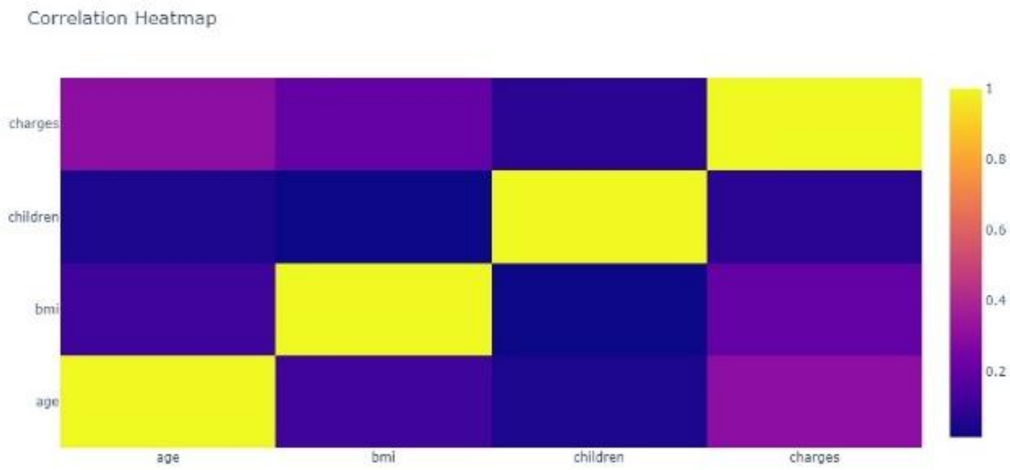

Figure 6. Correlation Heatmap

Correlation denotes an association, serving as a metric to gauge the relationship between two variables. Positive correlation indicates that as one variable incerases or decreases, the other variable follows suit in the same direction. When two variables move in tandem, they are positively correlated, with a perfect positive correlation represented by a value of '1'. Conversely, negative correlation arises when one variable's increase corresponds to the decrease of the other variable, and vice versa. This indicates a negative correlation, with a value of '-1' indicating a perfect negative correlation. A value of '0' suggests no correlation between the variables.

A Correlation Heatmap is shown in Figure 6, which proves that none of the variables in our dataset are highly correlated with each other.

5.4. Modeling

Before applying clustering algorithms and statistical analysis, the dataset comprising BMI-related health service costs underwent preprocessing. This step involved handling missing values, normalization, and feature selection to ensure data quality and relevance.

The results from K-means and agglomerative clustering methods provided distinct perspectives on the segmentation of health service costs related to varying BMI levels. The subsequent ANOVA analysis allowed for rigorous statistical validation of the cost differences among these clusters.

The combined application of K-means clustering, agglomerative clustering, and ANOVA facilitated a comprehensive examination of BMI-related financial burdens on health services. This multifaceted approach allowed for the identification of distinct cost clusters, their hierarchical relationships, and the statistical validation of disparities in healthcare expenses attributed to different BMI categories.

The insights gained from this modeling process offer valuable information for healthcare policymakers, enabling them to strategize and allocate resources effectively to address the financial implications of varying BMI levels on health services.

5.4.1. K – Means Clustering

K-means clustering was utilized to categorize the health service costs associated with different BMI groups. This unsupervised machine learning technique partitions the data into distinct clusters based on similarities in cost patterns. The aim was to identify clusters representing varying degrees of BMI-related expenses incurred by health services.

Cluster analysis, also known as clustering, involves grouping a collection of objects in a manner where objects within the same cluster share more similarities with each other than with those in other clusters. The degree of similarity between two points is established by their distance from one another. K-means clustering aims to minimize distances within clusters while maximizing the separation between different clusters. Essentially, the K-means algorithm identifies 'k' number of centroids and then assigns each data point to the closest cluster, striving to keep the centroids as compact as possible.

The 'means' in the K-means refers to averaging of the data, that is, finding the centroid.

5.4.2. Agglomerative Clustering

In addition to K-means, agglomerative clustering was employed to further validate the clustering results. This hierarchical clustering method merges similar clusters iteratively, forming a dendrogram that assists in identifying optimal cluster formations within the dataset. The agglomerative approach provided insights into the structure of BMI-related cost groups and their interrelationships.

Agglomerative Clustering is a type of hierarchical clustering algorithm. It is an unsupervised machine learning technique that divides the population into several clusters such that data points in the same cluster are more similar and data points in different clusters are dissimilar.

Agglomerative clustering is the most common type of hierarchical clustering used to group objects in clusters based on their similarity. The algorithm starts by treating each object as a singleton cluster.

5.4.3. ANOVA (Analysis of Variance)

Following the clustering analysis, ANOVA was conducted to assess the statistical significance of cost variations among the identified clusters. This parametric test evaluated whether the mean differences in health service expenses across different BMI groups were significant or occurred by chance. ANOVA helped to determine if the cost disparities observed between the clusters were statistically meaningful.

5.5. Proposed Method

In contemporary research, the amalgamation of diverse analytical tools and programming languages offers an expansive landscape for investigating complex phenomena. This proposed method delineates the strategic fusion of Python and R libraries to conduct K-means clustering, agglomerative clustering, and ANOVA analysis in scrutinizing the financial implications of Body Mass Index (BMI) on health service costs.

The method commences with comprehensive data acquisition, ensuring the collection of robust datasets encompassing BMI-related health service costs. Python's pandas and R's tidyverse packages are utilized for initial data cleaning, handling missing values, and formatting adjustments to ensure data integrity and relevance.

Following data preparation, exploratory data analysis is conducted. Descriptive statistics aided by Python's pandas and R's summary functions facilitate an understanding of the distribution of health service costs across various BMI groups. This phase is augmented by data visualizations, employing Python's Matplotlib, Seaborn, and R's ggplot2 for generating insightful graphs and charts. Leveraging Python's scikit-learn library, K-means clustering is implemented. The Elbow Method assists in determining the optimal number of clusters representing distinct BMI- related cost groups. Subsequently, the K-means algorithm partitions the dataset, categorizing health service costs based on varying BMI levels.

Complementing the K-means results, agglomerative clustering is performed using R's "stats" package. This hierarchical clustering method creates a dendrogram, elucidating the clustering hierarchy and validating the results obtained from K-means clustering.

ANOVA analysis is conducted using R's built-in functions and the "stats" package. This statistical test assesses the significance of health service cost variations among the identified clusters derived from both K-means and agglomerative clustering. The outcomes provide crucial insights into the statistical significance of disparities in healthcare expenses across different BMI categories.

In essence, this method amalgamates the strengths of Python and R, harnessing their libraries and functionalities, to unravel intricate patterns in health service costs associated with varying BMI levels, thereby paving the way for informed decision-making in healthcare management.

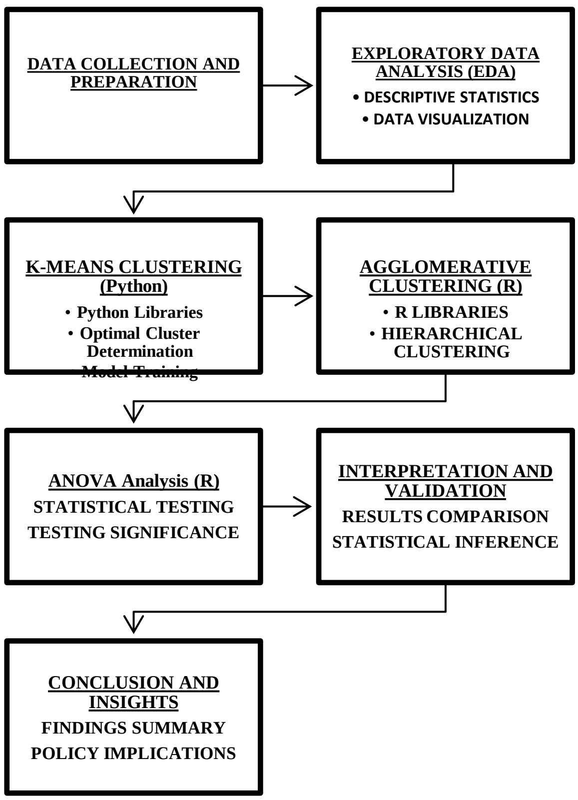

Figure 7. Proposed Method

6. Experimental Results

At times, caution is necessary when relying on the means due to its susceptibility to be influenced by outliers. Outliers or extreme cases can significantly impact the mean, prompting us to adopt two measures for graphically representing Charge across Age Categories: the mean and the median. The median, situated at the middle of a sorted list of numbers, whether ascending or descending, often provides a more descriptive representation of the dataset than the average. It serves as an alternative to the mean, especially in instances where outliers exist within the sequence, potentially skewing the average value.

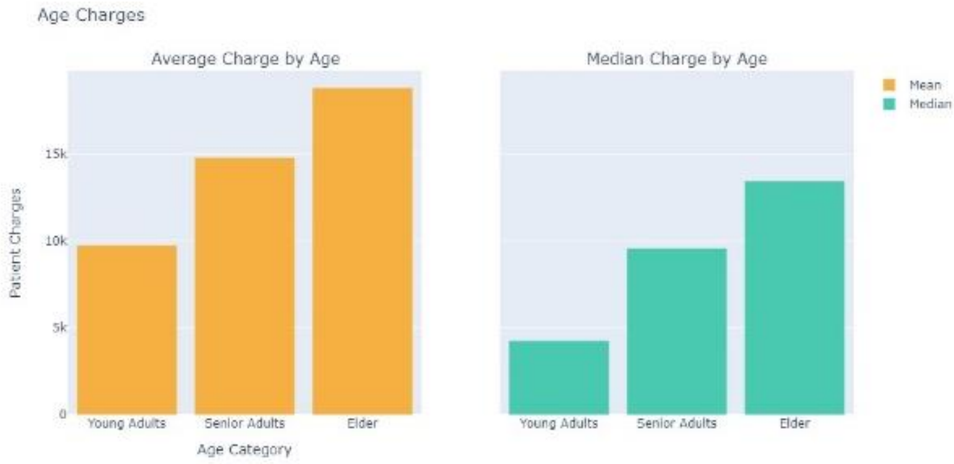

Figure 9. AGE variable charges

The mean medical care expenditure for the elderly stood at USD 18,026 per person, in contrast to USD 14,987 for the intermediate age group and USD 9,875 for the younger category (Figure 9). This indicates a higher average healthcare bill among the elder population compared to both the intermediate and younger age groups.

Approximately fifty percent of the total public expenditure on personal health care was allocated to the elderly population. Among this demographic, Medicare payments accounted for 44 percent of their personal health care expenses, while an additional 13 percent was covered by Medicaid. In contrast, private financing, mainly through private health insurance and direct payments, emerged as the primary payment channel for younger age groups, covering approximately 70 percent of their healthcare expenditures.

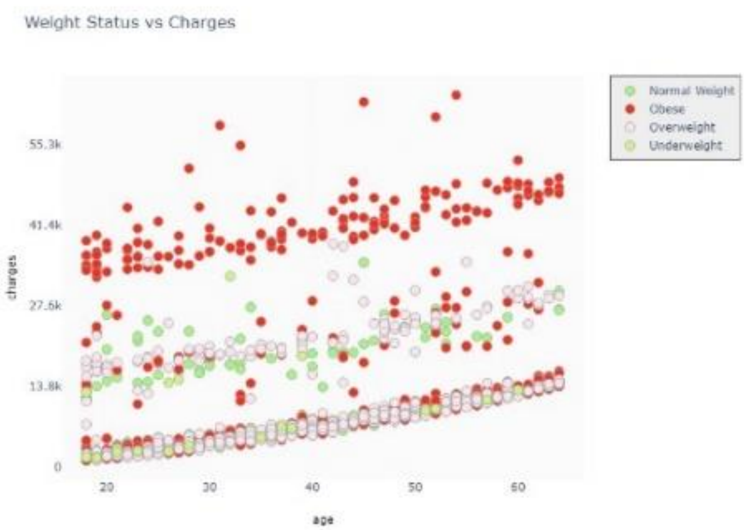

Figure 10. Weight vs Charge status

The WHO (World Health Organization) defines categories using the cut-off points: an individual with a BMI between 25.0 and 30.0 is considered to be ‘overweight’; a BMI greater than 30.0 is defined as ‘obese’.

Obesity stands out as one of the most significant health concerns worldwide, transitioning from a problem primarily observed in affluent nations to one that affects individuals across all income brackets. From the data (Figure 10), it's evident that the obese category incurs the highest healthcare expenditure, followed by the overweight group, with those of normal weight ranking lower. Notably, the underweight category reflects the lowest healthcare spending. Another noteworthy observation is the consistent trend of increased healthcare consumption among patients as they age, irrespective of their weight status.

The upcoming graph (Figure 11) will distinctly illustrate the differentiation between obese individuals who smoke and those who do not.

Figure 11. Obese smokers and non-smokers

Smoking stands as one of the most significant global health crises, contributing to the ill health of millions of individuals. Researchers estimate that approximately 8 million people suffer premature deaths annually due to smoking-related causes, resulting in a substantial burden on public health. This issue has persisted as a major health concern for numerous decades. Over the entire span of the 20th century, it's estimated that nearly 100 million people experienced premature mortality attributed to smoking, with the majority of these cases occurring in affluent nations.

Figure 12. Obese smokers and non-smokers

An evident demarcation is noticeable (Figure 12) in the patient charges among obese individuals who smoke versus those who do not. It's apparent that obese patients who are active smokers incur expenses nearly four times higher than their non-smoking counterparts within the same category. On average, smokers spend upwards of USD 35,000, while non-smokers spend approximately USD 8,000, with variations depending on age.

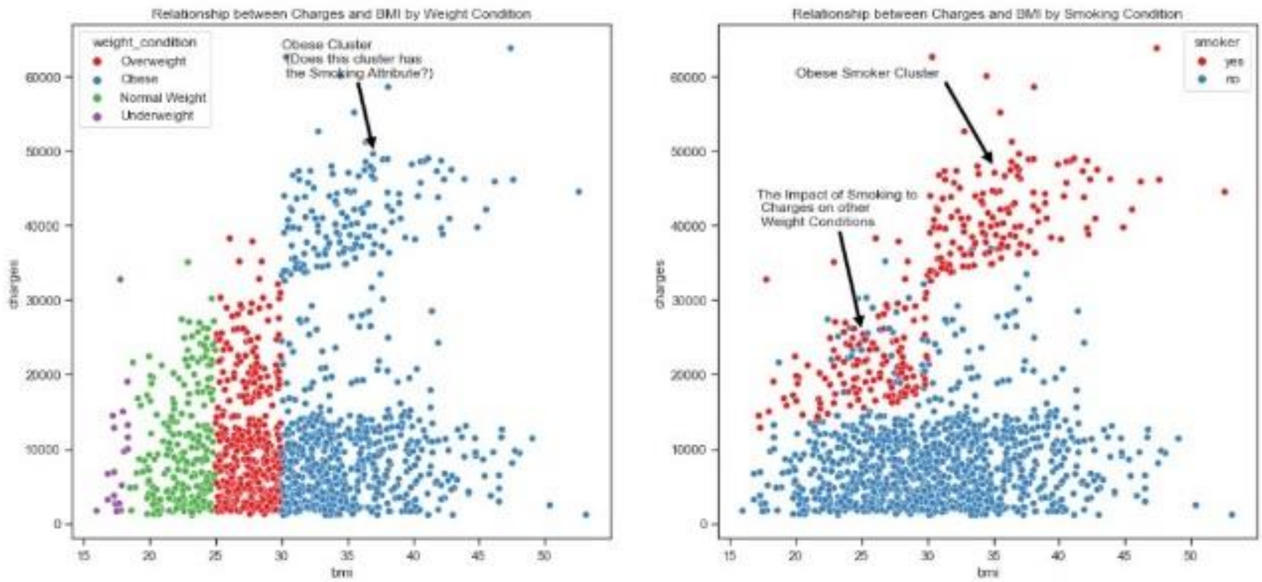

Figure 13. Relationships

Figure 13 provides a comprehensive depiction of the disparity in patient consumption segregated by weight condition. It's evident that the greatest consumption is attributed to patients classified as Obese, followed by those in the Overweight, Normal Weight, and Underweight categories, in descending order. The graph unmistakably reveals that a patient's weight condition significantly influences their incurred charges.

The following chart illustrates the correlation between charges based on smoking condition. A notable observation from both charts is the distinct visibility of the cluster comprising obese smokers. Meanwhile, the remaining clusters depict the discernible impact of smoking on individuals across various weight conditions. Specifically, the graph highlights how smoking influences healthcare charges within different weight categories, with the obese smoker cluster standing out prominently.

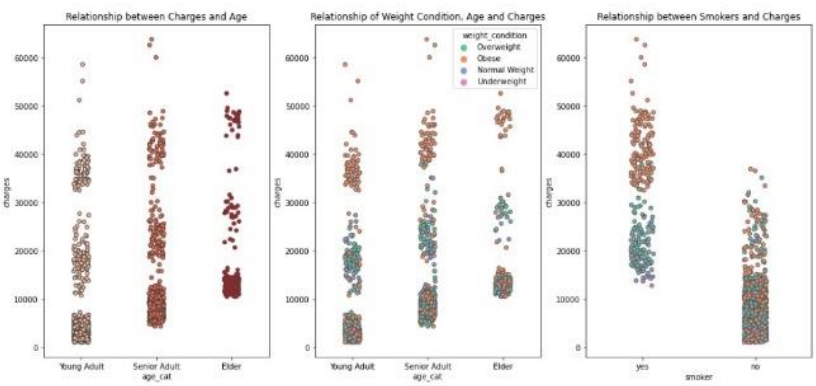

Figure 14. Relationships

Figure 14 displays the integrated relationships explained thus far, effectively showcasing the disparities among the mentioned categories in a single comprehensive representation. This chart conveniently summarizes the observed differences, offering a consolidated view of the various relationships between the mentioned categories.

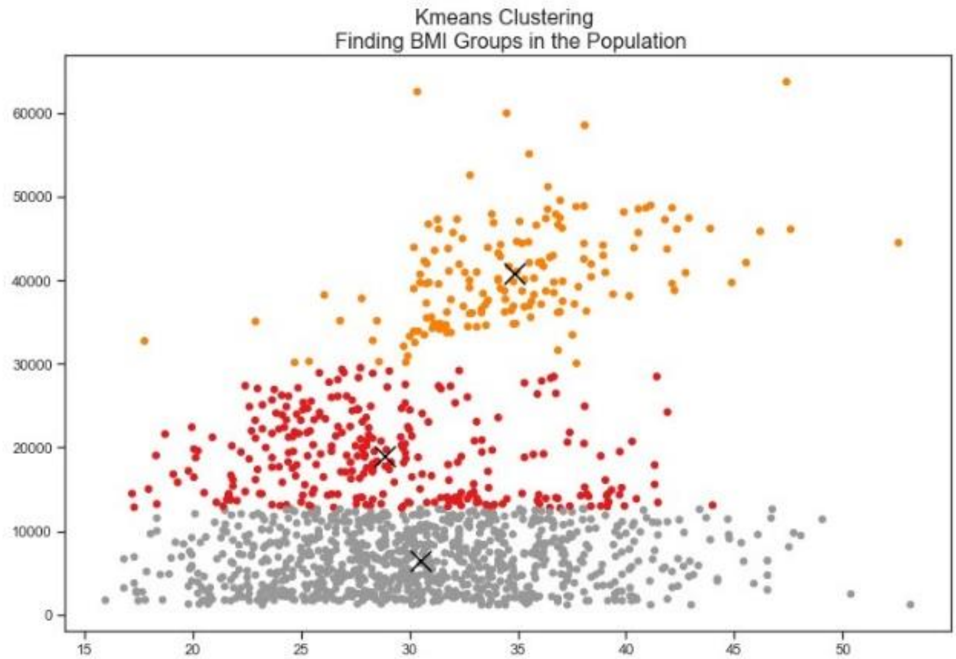

Figure 15. K means clustering

The chart depicted in Figure 15 displays a total of three clusters, visually represented by distinct colors and marked with their respective centroids. Each cluster is distinctly discernible within the visualization, aiding in the clear identification of their individual traits and characteristics.

The K-means clustering graph partitions clear groups according to BMI index and healthcare service expenditures, providing valuable insights into diverse spending patterns across different BMI categories. This interpretation assists in comprehending the unique characteristics of each cluster, thereby offering implications for healthcare management strategies and the formulation of policies.

The categorization of patient charges for health services into three distinct groups sheds light on varying healthcare spending patterns. Firstly, among these groups, a significant portion consists of patients who incur substantial healthcare expenses. This expenditure is often attributed to poor lifestyle habits that negatively affect their health. These individuals may require more extensive medical attention, treatments, or interventions due to their habits, leading to higher healthcare costs.

Secondly, another segment encompasses patients with moderate spending on health services. This group possibly comprises individuals who lack well-established health habits or preventive measures, resulting in a need for considerable healthcare support. Although they do not spend as much as the first group, their healthcare expenses still indicate a requirement for medical care and services.

Lastly, there exists a distinct cohort characterized by patients exhibiting excellent health habits. These individuals demonstrate notably lower average spending on health services compared to the other groups. Their emphasis on healthy living, proactive healthcare measures, and perhaps a lifestyle focused on wellness and preventive care contribute to reduced healthcare expenses.

These three delineated groups offer valuable insights into the diverse factors influencing healthcare spending, ranging from lifestyle choices and habits to the level of emphasis placed on preventive healthcare measures. Understanding these distinctions can aid in tailoring healthcare interventions, public health initiatives, and personalized approaches to better meet the needs of distinct patient groups.

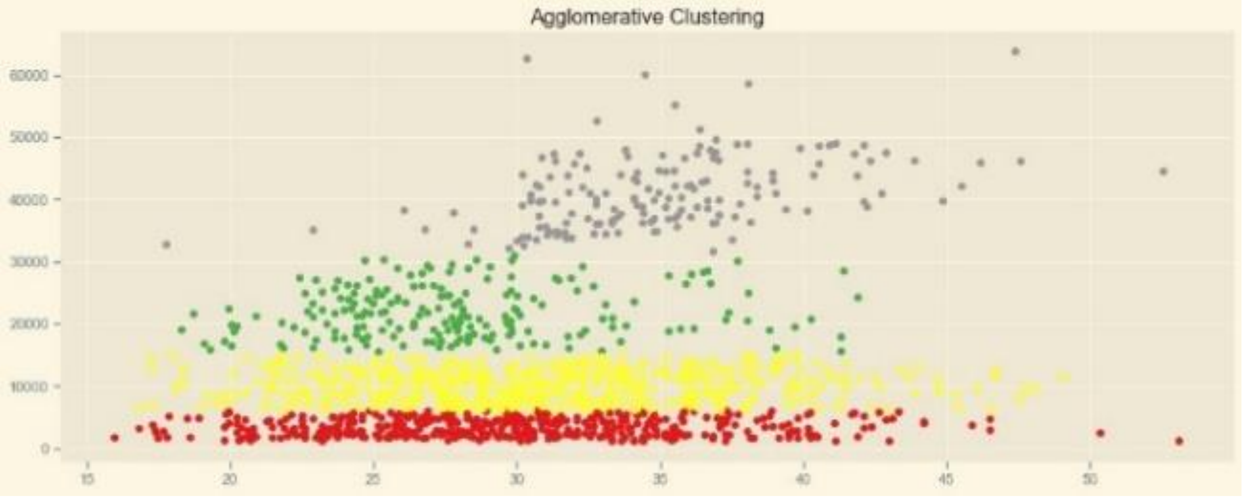

Figure 16. Agglomerative clustering

According to Figure 16, the presence of four clusters is noticeable.

Cluster 1 - "Minimal Service Utilization": This cluster represents individuals with a lower BMI range. Health service consumption in this group indicates minimal healthcare utilization. They require fewer medical interventions, visits, or treatments related to BMI.

Cluster 2 - "Regular Preventive Care": Individuals falling within an average BMI range belong to this cluster. Health service consumption shows regular or periodic preventive care visits, routine screenings, and moderate healthcare utilization related to managing BMI.

Cluster 3 - "Increased Medical Interventions": Individuals in this cluster have a higher BMI, indicating increased health risks. Health service consumption demonstrates higher medical interventions, frequent doctor visits, medication prescriptions, and potential hospitalizations due to BMI-related issues.

Cluster 4 - "Intensive Healthcare Utilization": This cluster consists of individuals with significantly high BMIs, indicating severe health risks. Health service consumption in this group signify extensive medical care, specialized treatments, frequent hospital stays, surgeries, or chronic condition management related to high BMI levels.

7. Discussion

The financial burden imposed by Body Mass Index (BMI) on health services has been a subject of increasing concern and interest. Through a comparative cost assessment utilizing big data, this study aims to shed light on the significant impact of BMI on healthcare expenditure. The discussion below outlines key findings and implications derived from the comprehensive analysis. The analysis of extensive datasets has revealed a substantial correlation between BMI and healthcare costs. Individuals categorized as overweight or obese exhibit a notably higher utilization of healthcare resources compared to those within healthier BMI ranges. The increase in medical expenses spans various aspects of healthcare services, encompassing primary care visits, hospital admissions, pharmaceutical expenditures, and interventions related to obesity-related comorbidities. The findings underscore the cascading effect of BMI on healthcare spending, highlighting a compelling need for strategic interventions and preventive measures. Notably, the disproportionately higher costs associated with obese individuals emphasize the urgency of proactive initiatives aimed at curbing obesity rates within populations. Investing in preventive healthcare strategies, lifestyle modification programs, and early interventions could potentially mitigate the escalating healthcare costs attributed to elevated BMI.

Additionally, this comparative cost assessment serves as a catalyst for ongoing research and policy formulation aimed at addressing the systemic challenges posed by BMI-related healthcare costs. Collaborative efforts involving healthcare stakeholders, policymakers, and technology innovators are imperative to develop innovative solutions, improve healthcare delivery models, and implement evidence-based policies to alleviate the burgeoning financial strain.

8. Conclusion

A healthy lifestyle is pivotal in maintaining fitness, vitality, and lowering the risk of various diseases. As per the World Health Organization (WHO), healthy living is a holistic approach that enables individuals to experience a fuller spectrum of life's facets. It involves practices that reduce the likelihood of severe illness or premature mortality. Health encompasses not only the absence of disease but also encompasses physical, mental, and social well-being. Effectively managing stress in constructive ways, rather than resorting to habits like smoking or excessive alcohol consumption, contributes to reducing the cellular wear and tear on one's body. Adopting a healthy lifestyle not only aids in longevity but also enhances the quality and comfort of life. The landscape of healthcare pricing remains opaque, leaving both patients and clinicians unaware of the true costs. With a surge in men enrolling in high-deductible health plans, there's a growing focus on the affordability of healthcare services.

Smoking remains among the most significant global health challenges, adversely affecting the well-being of millions worldwide. It's estimated that each year, approximately 8 million people succumb to premature deaths due to smoking-related causes, leading to a substantial population living in poor health. This issue has persisted as a major health concern for numerous decades. Throughout the 20th century, an estimated 100 million individuals faced premature mortality due to smoking, predominantly in affluent countries.

Obesity, likewise, stands as a monumental global health issue that has transcended its origins from being solely prevalent in affluent nations to encompassing all income levels. As per the Global Burden of Disease study, in 2017 alone, 4.7 million individuals met premature deaths attributable to obesity-related factors.

This paper provides an in-depth analysis of the multitude of factors that exert influence on healthcare expenditure. Our comprehensive findings underscore the pivotal role of adopting a healthier lifestyle, integrating consistent exercise routines, and refraining from smoking, among various other determinants, in considerably curbing healthcare expenses. The data illustrates a strong correlation: individuals demonstrating a propensity towards an unhealthy lifestyle, marked by elevated Body Mass Index (BMI) and habitual smoking, tend to incur notably higher costs for health services. Furthermore, our research consistently highlights the association between advancing age and the amplification of medical expenditures. The evidence presented emphasizes the critical impact of lifestyle choices and health behaviors on healthcare spending trends. Those embracing healthier habits exhibit a tangible reduction in healthcare costs, whereas those with unhealthy lifestyle tendencies face heightened financial burdens due to increased healthcare needs. Additionally, our findings underline the inevitable escalation of medical costs with advancing age, shedding light on the importance of proactive healthcare measures and interventions across different age groups.

Upon experimental results, we've identified three distinct patient profiles based on their healthcare expenditure. Our research highlights the stark contrast in spending patterns among these patient groups, emphasizing the strong association between lifestyle choices and healthcare expenses. The first group's higher healthcare costs underscore the impact of unhealthy behaviors on medical spending, while the third group's lower expenses align with their proactive approach to health management. Understanding these distinctions could aid in tailoring interventions and promoting healthier lifestyles to mitigate healthcare costs.

Both smoking and being overweight are substantial threats to individual health, each contributing significantly to an array of health complications across diverse physiological and psychological domains. Understanding the multifaceted health implications of these habits is crucial in shaping effective interventions, healthcare policies, and public health campaigns aimed at mitigating the extensive consequences of smoking and obesity on individual well-being. Exploring these interconnected health risks provides a critical foundation for advocating proactive measures to promote healthier lifestyles and enhance global health outcomes.

9. Future Work

As part of our research paperwork, we utilized comparative cost assessment to unravel the multifaceted dimensions of the financial toll imposed by varying BMI levels on health services. By analyzing and comparing healthcare expenditure data from diverse demographic groups, this study aims to uncover the differential impact of BMI on healthcare costs. Moreover, it aims to explore preventive measures, interventions, and policy frameworks that can potentially mitigate the economic strain on healthcare systems resulting from BMI-related health complications.

Utilize advanced data analytics and predictive modeling techniques to anticipate the future economic consequences of BMI on healthcare services. Employ machine learning algorithms and predictive analytics to forecast patterns in healthcare expenditure and identify high-risk populations, facilitating the creation of precise interventions. Explore the effectiveness and cost-efficiency of different intervention strategies designed to decrease BMI and subsequently mitigate the impact on healthcare spending. Assess the effectiveness of preventive measures, lifestyle interventions, community health programs, and policy initiatives in their capacity to mitigate BMI-related healthcare expenses. Conduct comparative cost evaluations across diverse global regions to explore the variation in healthcare expenditures linked to BMI. Investigate the impact of cultural, economic, and healthcare system disparities on the financial burden of BMI-related healthcare costs, fostering a comprehensive global understanding of this issue. Explore the impact of social determinants of health (e.g., income, education, access to healthcare) on BMI-related healthcare expenditures. Analyze how these factors interact with BMI to influence healthcare costs and develop strategies that address social determinants to alleviate the financial burden on health services.

Investigate the role of technological advancements, such as telemedicine, wearable devices, and digital health solutions, in managing BMI-related health conditions, and assess the potential of technology-driven interventions to optimize healthcare delivery and reduce associated costs.

By exploring these avenues for future research, it will be possible to gain a more nuanced understanding of the evolving relationship between BMI and healthcare expenditures, leading to the development of targeted interventions and policy frameworks aimed at mitigating the financial impact on health services while promoting improved health outcomes.

References

[1]. K. F. Adams *et al.*, “Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old,” *N. Engl. J. Med.*, vol. 355, no. 8, pp. 763–778, 2006.

[2]. C. M. Apovian and N. Gokce, “Obesity and cardiovascular disease,” *Circulation*, vol. 125, no. 9, pp. 1178–1182, 2012.

[3]. D. E. Arterburn, & A. P. Courcoulas, “Bariatric surgery for obesity and metabolic conditions in adults”. *BMJ*, no. 349, page 3961, 2014.

[4]. K. Bhaskaran, I. Dos-Santos-Silva, D. A. Leon, I. J. Douglas, and L. Smeeth, “Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3•6 million adults in the UK,” *The Lancet Diabetes & Endocrinology*, vol. 6, pp. 944–953, 2018.

[5]. J. Cawley and C. Meyerhoefer, “The medical care costs of obesity: an instrumental variables approach,” *J. Health Econ.*, vol. 31, no. 1, pp. 219–230, 2012.

[6]. J. Cawley and C. Meyerhoefer, “The medical care costs of obesity: an instrumental variables approach,” *J. Health Econ.*, vol. 31, no. 1, pp. 219–230, 2012.

[7]. S. H. Chang, C. R. Stoll, Y. M. Tzeng, and D. A. Freedman, “Predictors of hospital charges in bariatric surgery: the role of patient comorbidity,” *Obesity Surgery*, vol. 25, no. 6, pp. 1108–1114, 2015.

[8]. Y. C. Chooi, C. Ding, & F. Magkos, “The epidemiology of obesity” *Metabolism*, no. 92, page. 6-10, 2019.

[9]. S. Colagiuri *et al.*, “The cost of overweight and obesity in Australia,” *Med. J. Aust.*, vol. 192, no. 5, pp. 260–264, 2010.

[10]. G. Danaei *et al.*, “The preventable causes of death in the United States: Comparative risk assessment of dietary, lifestyle, and metabolic risk factors,” *PLoS Med.*, vol. 6, no. 4, p. e1000058, 2009.

[11]. J. B. Dixon, “The effect of obesity on health outcomes,” *Mol. Cell. Endocrinol.*, vol. 316, no. 2, pp. 104–108, 2010.

[12]. E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, “Annual medical spending attributable to obesity: payer-and service-specific estimates: Amid calls for health reform, real cost savings are more likely to be achieved through reducing obesity and related risk factors,” *Health Aff. (Millwood)*, vol. 28, no. 5, pp. w822-31, 2009.

[13]. K. M. Flegal, B. I. Graubard, D. F. Williamson, & M. H. Gail, M. H. “Excess deaths associated with underweight, overweight, and obesity”. *JAMA*, vol. 293, no. 15, pp. 1861-1867, 2005.

[14]. K. M. Flegal, B. K. Kit, H. Orpana, & B. I. Graubard, “Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis”. *JAMA*, vol. 309, no. 1, pp. 71-82, 2013.

[15]. K. R. Fontaine, D. T. Redden, C. Wang, A. O. Westfall, & D. B. Allison, “Years of life lost due to obesity”. *JAMA*, vol. 289, no. 2, pp. 187-193, 2003.

[16]. E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, “Annual medical spending attributable to obesity: payer-and service-specific estimates: Amid calls for health reform, real cost savings are more likely to be achieved through reducing obesity and related risk factors,” *Health Aff. (Millwood)*, vol. 28, no. 5, pp. w822-31, 2009.

[17]. K. M. Flegal, B. I. Graubard, D. F. Williamson, & M. H. Gail, “Excess deaths associated with underweight, overweight, and obesity”. *JAMA*, vol. 293, no. 15, pp. 1861-1867, 2005.

[18]. K. M. Flegal, B. K. Kit, H. Orpana, & B. I. Graubard, “Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis”. *JAMA*, vol. 309, no. 1, pp. 71-82. 2013.

[19]. K. R. Fontaine, D. T. Redden, C. Wang, A. O. Westfall, & D. B. Allison, “Years of life lost due to obesity”. *JAMA*, vol. 289, no. 2, pp. 187-193, 2003

[20]. K. M. Flegal, M. D. Carroll, B. K. Kit, & C. L. Ogden, (2012). “Prevalence of obesity and trends in the distribution of body mass index among US adults,1999-2010.” *JAMA*, vol. 307, no. 5, pp. 491-497, 2012.

[21]. K. M. Flegal, D. Kruszon-Moran, M. D. Carroll, C. D. Fryar, & C. L. Ogden, “Trends in obesity among adults in the United States, 2005 to 2014,” *JAMA*, vol. 315, no. 21, pp. 2284-2291, 2016.

[22]. E. S. Ford and A. H. Mokdad, “Epidemiology of obesity in the Western Hemisphere,” *J. Clin. Endocrinol. Metab.*, vol. 93, no. 11 Suppl 1, pp. S1-8, 2008.

[23]. W. T. Garvey *et al.*, “American Association of Clinical Endocrinologists and American College of Endocrinology comprehensive clinical practice guidelines for medical care of patients with obesity,” *Endocrine Practice*, vol. 22, no. suppl_3, pp. 1–203, 2016.

[24]. C. M. Hales, M. D. Carroll, C. D. Fryar, and C. L. Ogden, “Prevalence of obesity among adults and youth: United States, 2015-2016,” *NCHS Data Brief*, no. 288, pp. 1–8, 2017.

[25]. M. Hamer and E. Stamatakis, “Overweight and obesity as potential determinants of psychosocial health among children and adolescents,” *Journal of Obesity*, pp. 1–7, 2008.

[26]. M. Hasan, I. Sutradhar, T. Akter, R. Das Gupta, and H. Joshi, “Economic burden of overweight and obesity: empirical estimates and projections in the WHO African Region,” *BMC Public Health*, vol. 19, no. 1, pp. 1–13, 2019.

[27]. S. Herpertz and M. De Zwaan, “Obesity and eating disorders,” *Obesity Medicine*, vol. 15, pp. 1–5, 2019.

[28]. S. B. Heymsfield, T. A. Wadden, & J. I. Mechanick, “Obesity: evaluation and treatment essentials.” *CRC Press*, 2017.

[29]. J. O. Hill, H. R. Wyatt, G. W. Reed, & J. C. Peters, “Obesity and the environment: where do we go from here?.” *Science*, vol. 299, no. 5608, pp. 853-855, 2003.

[30]. F. B. Hu, “Obesity epidemiology.” *Oxford University Press*, 2008.

[31]. M. D. Jensen, D. H. Ryan, C. M. Apovian, J. D. Ard, A. G. Comuzzie, K. A. Donato, & S. Z. Yanovski, “2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults.” *Circulation*, vol. 129, no. 25_suppl_2, pp. S102-S138, 2013.

[32]. S. Kahan, & J. E. Manson, “Obesity treatment, beyond the guidelines: practical suggestions for clinical practice.” *JAMA*, vol. 317, no. 1, pp. 27-28, 2017.

[33]. T. Kelly, W. Yang, C.-S. Chen, K. Reynolds, and J. He, “Global burden of obesity in 2005 and projections to 2030,” *Int. J. Obes. (Lond)*, vol. 32, no. 9, pp. 1431–1437, 2008.

[34]. P. G. Kopelman, “Obesity as a medical problem.” *Nature*, vol. 404, no. 6778, pp. 635-643, 2007.

[35]. R. F. Kushner and D. H. Bessesen, *Treatment of the obese patient*. Elsevier Health Sciences, 2007.

[36]. I. Kyrou, & C. Tsigos, “Obesity in the post-COVID-19 world: a concise review of literature and a call for action.” *Obesity*, vol. 28, no. 11, pp. 2002-2009, 2018.

[37]. F. S. Luppino *et al.*, “Over-weight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies,” *Archives of General Psychiatry*, vol. 67, no. 3, pp. 220–229, 2010.

[38]. A. H. Mokdad, J. S. Marks, D. F. Stroup, and J. L. Gerberding, “Actual causes of death in the United States, 2000,” *JAMA*, vol. 291, no. 10, pp. 1238–1245, 2004.

[39]. K. M. Narayan, J. P. Boyle, T. J. Thompson, E. W. Gregg, and D. F. Williamson, “Effect of BMI on lifetime risk for diabetes in the US,” *Diabetes Care*, vol. 30, no. 6, pp. 1562–1566, 2007.

[40]. M. Ng *et al.*, “Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013,” *Lancet*, vol. 384, no. 9945, pp. 766–781, 2014.