International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Multifactorial Risk Analysis and Prediction for Osteoporosis among Adults using Exploratory Data Analysis

Uma G[1], Sri Aishwarya A M[2]

[1]Associate Professor, Department of Statistics, PSG College of Arts & Science, Coimbatore - 641014, Tamil Nadu, India
Email: *uma[at]psgcas.ac.in*

[2]Post Graduate, Department of Statistics, PSG College of Arts & Science, Coimbatore - 641014, Tamil Nadu, India
Email: *sriaishwarya2703[at]gmail.com*

Abstract: ***Osteoporosis is a silent, progressive bone disease that significantly increases the risk of fractures and morbidity, particularly among aging populations. This study aims to analyse the multifactorial risk factors contributing to osteoporosis among adults using statistical methods. The dataset, sourced from Kaggle, comprises 1, 958 samples with variables spanning demographic (age, gender, race), lifestyle (physical activity, smoking, alcohol), nutritional (calcium and vitamin D intake), and medical (hormonal changes, medications, prior fractures) domains. Using Python and Power BI, the data is analysed with Exploratory Data Analysis (EDA), Chi - square test, and two - sample t - test to determine the descriptive measures, associations and patterns among the variables of the dataset. Results revealed age as the most significant predictor, with a higher prevalence observed in individuals aged 40 and above. Other contributing factors included underweight status, sedentary lifestyle, postmenopausal hormonal status, and corticosteroid medication use. However, variables such as family history and nutrition showed weak associations. This paper highlights the importance of early identification of at - risk individuals through basic demographic and lifestyle indicators to support preventive strategies.***

Keywords: Osteoporosis, Risk Factors, Descriptive Statistics, EDA, Chi - Square Test, t - Test, Bone Health, Lifestyle Analysis.

1. Introduction

In recent years, the term ‘Osteoporosis’ is commonly heard word from the public of both male and female population, but mostly in the age group of 50 and above especially among women after menopause due to the sudden decrease in estrogen, which usually protects against Osteoporosis. As per International Osteoporosis Foundation (IOF) survey, 1 in 3 women and 1 in 5 men will experience fractures due to Osteoporosis once the people reach the age of 50 years and in recent years it is becoming an important public health related problem globally. Osteoporosis is a multifactorial metabolic bone disorder characterized by reduced Bone Mineral Density (BMD) and structural deterioration of bone tissue, leading to an increased risk of fragility fractures. Often referred to as the “silent disease, ” osteoporosis progresses without symptoms until fractures occur, commonly affecting the hip, spine, and wrist. With the global rise in aging populations, sedentary lifestyles, and poor nutritional habits, osteoporosis has emerged as a significant public health concern, contributing to long - term disability and elevated healthcare costs. This condition is influenced by a wide range of interrelated factors, including age, gender, hormonal changes, genetic predisposition, physical inactivity, inadequate calcium and vitamin D intake, body weight, smoking, alcohol consumption, and medication history. Postmenopausal women, in particular, are at elevated risk due to estrogen deficiency, which accelerates bone resorption. While clinical tools like DEXA scans and the FRAX model assist in diagnosing and evaluating fracture risk, they are often inaccessible or underutilized in resource - limited settings. Therefore, understanding the distribution and impact of common demographic, lifestyle, and medical variables on osteoporosis is crucial for developing early screening strategies and population - level interventions. This paper aims to conduct a comprehensive descriptive and statistical analysis of these risk factors using a real - world dataset of adult individuals taken as secondary data from Kaggale. Through frequency distributions, cross - tabulations, and hypothesis testing, we identify key patterns and associations that can inform future predictive models and health policy planning.

2. Scope of the Study

Osteoporosis is a condition where bones become weak and fragile due to a loss of bone mineral density and changes in bone structure. Often called the "silent disease, " it usually affects the hip, spine, and wrist and shows no symptoms until a fracture occurs. It has become a growing health problem because of increasing elderly populations, inactive lifestyles, and poor nutrition. Many factors like age, gender, hormones, family history, physical activity, nutrition, body weight, smoking, alcohol, and certain medications can increase the risk of osteoporosis. This study aims to explore and describe how osteoporosis is distributed among different groups of people based on their demographic, lifestyle, and medical details. It uses visual tools and descriptive statistics to show patterns, trends, and relationships in the data. An interactive dashboard is also created to present the results clearly. The study’s findings will help in identifying people at higher risk and in supporting awareness and prevention strategies for better bone health.

Objectives of the Study

The aim of the study is to increase the knowledge about Osteoporosis – its causes and effects on the body

(pathophysiology), risk factors, as well as its prevention and treatment. The objective of the study is to gather evidence - based knowledge about osteoporosis and able to

- Predict the person under risk of developing Osteoporosis.
- Classify patients based on their bone health data
- Identify patterns and risk factors from patient's demographics
- Analyse lifestyle impacts on bone density.
- To create an interactive Power BI dashboard for visualizing key insights.

3. Review of the Study

This study was carried by various authors and considered for this study. Following are the some of the literatures which is relevant to this study are: Kanis et al. (2008) evaluated the global burden of osteoporosis and identified key risk factors such as increasing age, female gender, sedentary lifestyle, and poor nutrition. Their findings showed that over **30% of women and 20% of men above the age of 50** are affected by osteoporosis worldwide, emphasizing the need for early detection and public awareness. In another large - scale study, Looker et al. (2012) analyzed bone mineral density data from the femur neck and lumbar spine of U. S. adults and found that **10.3% of adults aged 50 and older** had osteoporosis at the femur neck, with significantly higher prevalence in women than men. This aligns closely with the demographic trends observed in the present study, particularly regarding age and gender. Additionally, Kanis et al. (2013) proposed European guidelines for osteoporosis management in postmenopausal women, highlighting crucial risk factors such as menopause, low calcium intake, physical inactivity, and family history. These results correlate with this study's stacked bar chart analysis, which showed higher osteoporosis risk among postmenopausal women and individuals with sedentary lifestyles or corticosteroid medication use.

4. Research Methodology

Data Source and Collection Method

The data for this study was sourced from the Kaggle platform, specifically from the Osteoporosis Dataset. This dataset includes 1958 samples from individuals with various demographic, lifestyle, and medical backgrounds, collected through surveys and questionnaires. The variables included in the dataset span age, gender, physical activity levels, smoking habits, alcohol consumption, calcium intake, family history of osteoporosis, previous fractures, and medical conditions. These observations were used to analyze the key risk factors contributing to osteoporosis. The data was obtained from secondary sources, ensuring that it is comprehensive and representative of different populations, allowing for detailed Exploratory Data Analysis (EDA) and Statistical techniques.

5. Results and Discussion

Table 1: Demographic Characteristics of Study Participants (N = 1958)

Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	992	50.66
	Female	966	49.34
Osteoporosis Status	Yes	979	50.00
	No	979	50.00
Age Group	18–30 years	654	33.39
	31–50 years	612	31.27
	51–70 years	478	24.41
	>70 years	214	10.93
Body Weight	Normal	1350	68.96
	Underweight	608	31.04
Calcium Intake	Sufficient	984	50.26
	Insufficient	974	49.74
Vitamin D Intake	Sufficient	967	49.39
	Insufficient	991	50.61
Physical Activity	Active	961	49.08
	Sedentary	997	50.92
Smoking	Yes	938	47.91
	No	1020	52.09
Alcohol Consumption	Moderate	961	49.08
	None	997	50.92
Hormonal Changes	Normal	988	50.46
	Menopausal/Postmenopausal	970	49.54
Medications	Corticosteroids	978	49.95
	None	980	50.05
Family History	Yes	990	50.56
	No	968	49.44
Prior Fracture	Yes	979	50.00
	No	979	50.00
Medical Conditions	Hyperthyroidism	678	34.63
	Rheumatoid Arthritis	633	32.33
	None	647	33.04

Table 2: Descriptive Statistics for Age of Study Participants

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Age	39.1	35	17.6	18	90

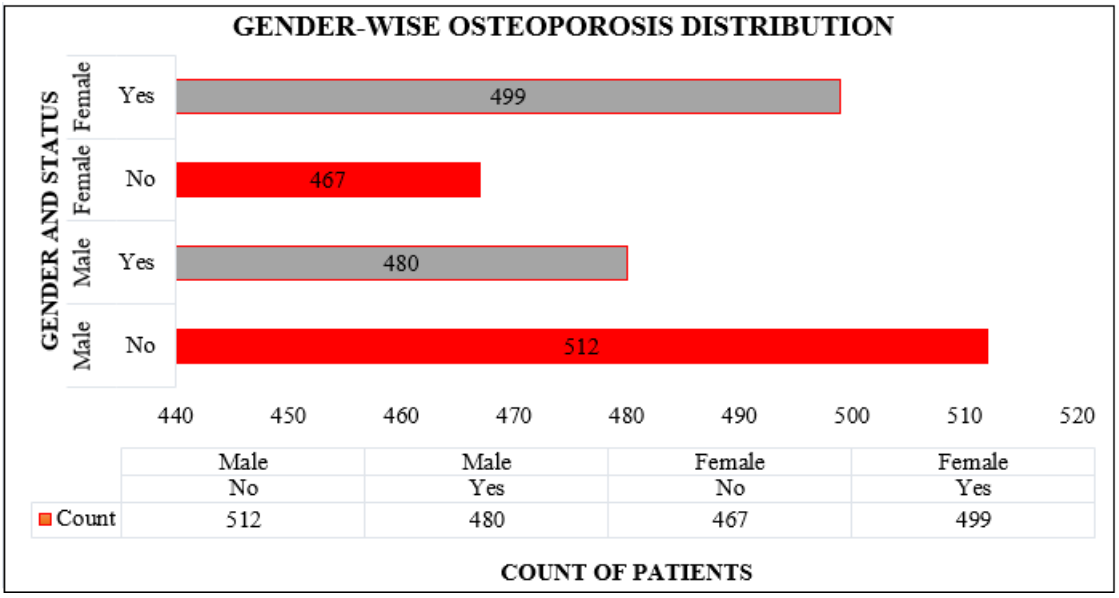

Figure 1: Bar Chart for the Target Variable Osteoporosis

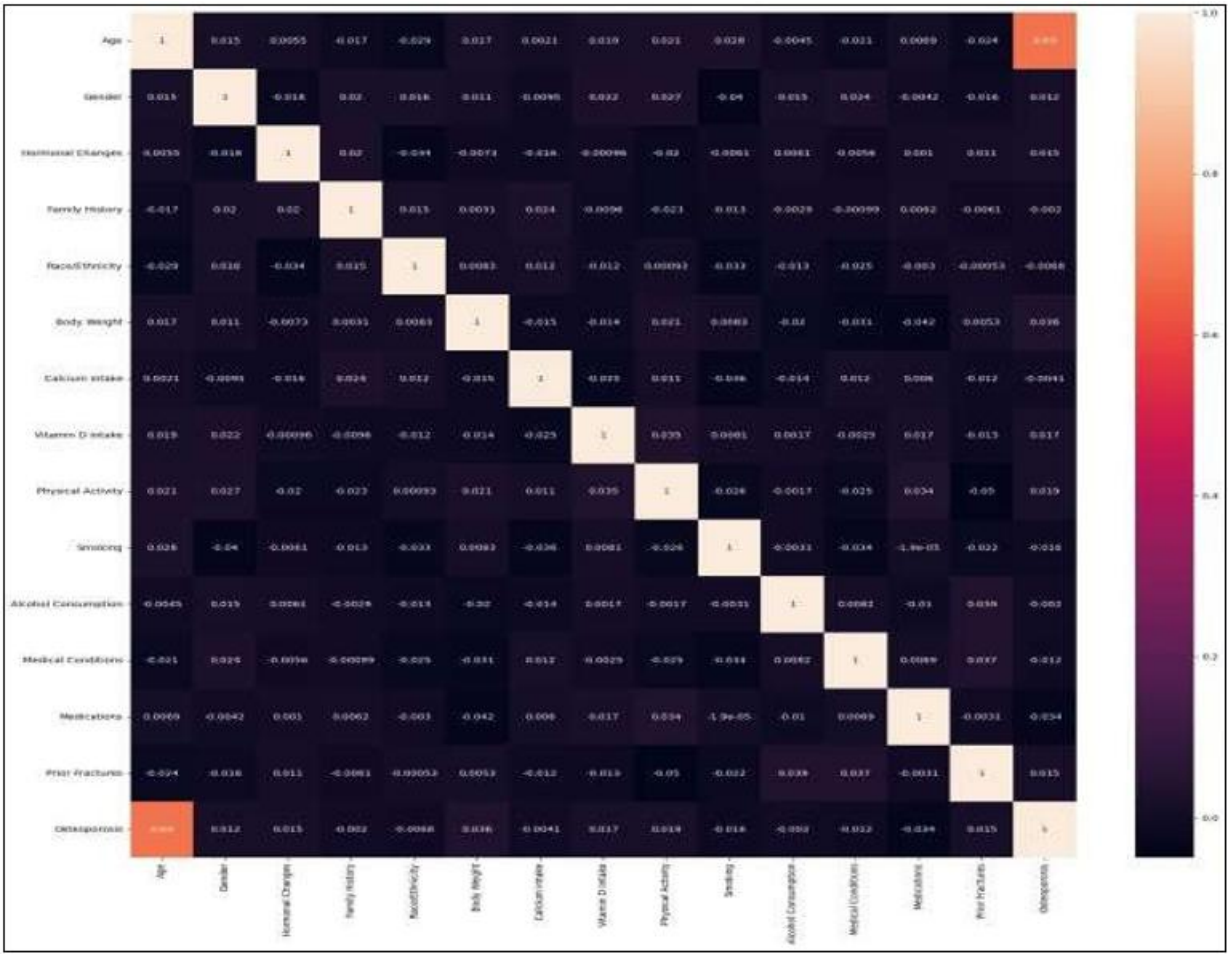

Figure 2: Heat map for the Variables

Results from Demographic Characteristics of Study Participants (Table 1)

- The Table 1 presents the demographic distribution of the 1958 participants included in the study. The sample is well - balanced in terms of gender, with 50.66% males (n = 992) and 49.34% females (n = 966). The prevalence of osteoporosis is evenly distributed, with 979 participants (50%) diagnosed with the condition and 979 (50%) without.

- In terms of age, a significant proportion of participants fall into the 18–30 age group (33.39%), followed by 31–50 years (31.27%), 51–70 years (24.41%), and only 10.93% are above 70 years**.** This wide age range supports the investigation of age - related risk factors for osteoporosis.
- For Body weight**,** 68.96% of participants have normal weight**,** while 31.04% are underweight, which may be a concern considering the impact of low body mass on bone density.
- When analyzing nutritional factors, calcium and vitamin D intake are almost evenly distributed, with around 50% having sufficient intake and 50% insufficient, providing a good contrast for evaluating dietary risk factors.
- Regarding lifestyle**,** 49.08% are physically active, and 50.92% are sedentary, while 47.91% are smokers, and 52.09% are non - smokers**.** Similarly, alcohol consumption is balanced, with 49.08% consuming alcohol moderately and 50.92% abstaining**.**
- In terms of hormonal changes, the sample includes 50.46% with normal hormonal levels and 49.54% menopausal/postmenopausal individuals**,** offering a comparative look at hormonal influence.
- For clinical risk factors, 49.95% of participants use corticosteroid medications, and 50.05% do not. Participants with a family history of osteoporosis make up 50.56%**,** while 49.44% do not have such a history**.** Similarly, prior fractures are equally split (50% yes, 50% no), suggesting a strong baseline for evaluating fracture risk.
- Finally, regarding medical conditions, the sample is fairly evenly distributed among participants with hyperthyroidism (34.63%), rheumatoid arthritis (32.33%), and no medical conditions (33.04%).

Results from Descriptive Statistics for Age of Study Participants (Table 2)

As shown in Table 2, the average age of the participants is 39.1 years**,** with most individuals around 35 years old**,** and an age spread ranging from 18 to 90 years. This wide age range reflects a diverse group of participants, which is important for studying osteoporosis. Since the risk of osteoporosis increases with age, having a good number of older adults in the sample adds strength and credibility to any conclusions drawn about age - related trends.

Results from Bar Chart for the Target Variable Osteoporosis (Figure 1)

In a sample of 1, 958 respondents, the gender distribution is nearly equal, with 992 males (50.66%) and 966 females (49.34%). Among males, 512 (51.6%) do not have osteoporosis, while 480 (48.4%) are affected. In females, 467 (48.34%) are not affected, and 499 (51.66%) have osteoporosis. Although both genders are nearly equally affected, the percentage of females with osteoporosis is slightly higher, indicating a marginally greater risk in females, consistent with clinical trends, especially post - menopause.

Results from Heat map for the Variables (Figure 2)

Heat map is the diagrammatic view on the correlation of multivariable in the data set especially huge data set with more variable. Accordingly shows that age (0.69) is the strongest predictor of osteoporosis, with the risk increasing significantly as age rises. Factors like calcium intake (0.014), vitamin D intake (0.017), and physical activity (0.019) shows weak positive correlations, suggesting minimal direct links to osteoporosis. Smoking (0.016) and alcohol consumption (- 0.004) have negligible correlations, indicating little direct effect, though they may indirectly influence bone health. Medical conditions (- 0.024), medications (- 0.0069), and prior fractures (- 0.024) also show weak negative correlations, meaning they do not strongly predict osteoporosis. While age is the primary risk factor, lifestyle, diet, and medical conditions still impact bone health through complex interactions not fully captured by this analysis.

Chi - Square Test and t - Test

From *Chi - square test results (χ^2 = 0.296, df = 2, p - value = 0.8626), there is no statistically significant association between medical conditions and osteoporosis. Since the p - value (0.8626)* is much greater than the conventional threshold of 0.05, the null hypothesis is accepted. This suggests that the occurrence of osteoporosis does not significantly differ across individuals with Hyperthyroidism, Rheumatoid Arthritis, or no medical condition.

The t - test shows a highly significant difference *in the average age between individuals with and without osteoporosis (t = - 42.306, df = 1956, p =0.0000).* The mean age for individuals without osteoporosis is 24.34 years, while for those with osteoporosis, it is 53.86 years. The 95% confidence interval for the difference in means is from - 30.88 to - 28.15, indicating that the true mean age of osteoporosis patients is significantly higher. Since the p - value is far below 0.05, reject the null hypothesis and conclude that age is significantly associated with osteoporosis, and individuals with the condition tend to be significantly older.

6. Conclusion

This study highlights age as the strongest predictor of Osteoporosis, with the risk significantly increasing as individuals' age. While lifestyle factors such as calcium intake, vitamin D intake, and physical activity shows weak positive correlation, their direct impact on osteoporosis remains minimal. Other factors like smoking, alcohol consumption, medical conditions, medications, and prior fractures have negligible or weak correlations, suggesting that these variables may influence bone health in more indirect and complex ways. Overall, the findings emphasize the crucial role of age in osteoporosis risk while acknowledging the multifactorial nature of the condition, influenced by lifestyle, diet, and medical history.

7. Recommendations

Based on the findings, it is recommended that interventions for Osteoporosis primarily focus on age - related risk factors, particularly for individuals approaching or in later stages of life. Additionally, promoting a healthy lifestyle that includes adequate Calcium and Vitamin D intake, as well as regular physical activity, could help mitigate osteoporosis risk. Although lifestyle and medical factors showed weak

correlations, should not be ignored, as their role in overall bone health may be more complex. Future research should explore the interactions between these variables in greater depth to better understand their indirect influence on Osteoporosis risk.

8. Limitations

- The study mainly looks at correlations, so it doesn't fully explain cause - and - effect relationships between the factors and osteoporosis.
- It doesn't take into account other potential influences, like genetic factors or environmental conditions, that could also affect bone health.
- The sample used may not reflect the diversity of the entire population, which means the results may not apply to everyone and changes according the geographical conditions.
- Some factors, like diet and medications, might have a bigger impact than the study suggests, but the weak correlations didn’t highlight this fully.

References

[1] **Kanis, J. A., Johnell, O., Oden, A., Johansson, H., & McCloskey, E. (2008).** Worldwide prevalence of osteoporosis and its impact on public health. Osteoporosis International, 19 (4), 399–408.

[2] **Looker, A. C., Frenk, S. M., & Hughes, J. P. (2012).** Osteoporosis and low bone mass in older U. S. adults: Data from the National Health and Nutrition Examination Survey, 2005–2010. NCHS Data Brief, (93), 1–8.

[3] **Kanis, J. A., McCloskey, E. V., Johansson, H., Cooper, C., Rizzoli, R., & Reginster, J. Y. (2013).** European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporosis International, 24 (1), 23–57.

[4] **Siris, E. S., Chen, Y. T., Abbott, T. A., Barrett - Connor, E., Miller, P. D., Wehren, L. E., & Berger, M. L. (2004).** Bone mineral density thresholds for pharmacological intervention to prevent fractures. Archives of Internal Medicine, 164 (10), 1108–1112.

[5] **Baim, S., & Krebs, J. (2008).** Bone mineral density testing in clinical practice. *Endocrinology and Metabolism Clinics of North America, 37* (1), 37 - 51.

[6] **Riggs, B. L., & Melton, L. J. (1995).** The worldwide problem of osteoporosis: Insights afforded by epidemiology. *Bone, 17* (3), 505 - 511.

[7] **Bauer, D. C., & Hochberg, M. C. (2000).** Epidemiology of osteoporosis. *Endocrinology and Metabolism Clinics of North America, 29* (2), 211 - 228.

[8] **Fuggle, N., & Cooper, C. (2014).** Osteoporosis and fractures: A global challenge. *Osteoporosis International, 25* (11), 2515 - 2522

[9] **Siris, E. S., et al. (2001).** Osteoporosis and fracture risk in men. *Bone, 29* (3), 200 - 208.

[10] **Lindsay, R., & Cosman, F. (2002).** The prevention and treatment of osteoporosis: Clinical practice guidelines. *American Journal of Medicine, 112* (2), 78 - 84.