Spearman's Rank Correlation and Factorial Analysis: Socio-Demographic Characteristics of Diabetic Hypertensive Patients Presenting to a Tertiary Care Hospital

Monika Kapoor[1,2,*]

[1]Department of Pharmacy Practice, ISF College of Pharmacy, Moga, Punjab, INDIA.
[2]Intern, Clinical Pharmacologist, Punjab Institute of Medical Science, Jalandhar, Punjab, INDIA.

ABSTRACT

Objectives: The main objective of this paper is to explore exploratory factor analysis and Spearman's Rank Correlation which indicates the relationship between the duration of the disease with socio-demographic characteristics of the patients presenting to the hospital. **Study Design:** Prospective Observational Study. **Methods:** The study was carried out in the Department of Medicine of the Punjab Institute of Medical Science and Hospital (PIMS), Jalandhar for a period of 6 months from June 2018 - November 2018 after Recieving approval from the Institutional Ethics Committee. After calculating the sample size using Epi Info, the study was conducted on a total of 250 patients. The study was analyzed using SPSS version 24. **Results:** The age group of the study was 22 to 86 years. According to the Kolmogorov Smirnov (KS) and Shapiro-Wilk (SW) p value was found to be $p = 0.009^a$ and $p = 0.058^b$ respectively. Age was found to be normally distributed because the p-value was found above the cut-off point with Skewness -0.200, Kurtosis 0.178 and standard error 0.762. It was observed that there is a significant moderate positive correlation between age and duration of HTN ($r = 0.408$, $p = 0.001$). Further analysis was observed that age, social habits and physical activity also has a wide impact on the duration of disease. **Conclusion:** Numerous factors are observed to be directly associated with the socio-demographic characteristics of the patients. During the study, it was observed that age, social habits and physical activity also has a wide impact on the duration of the disease itself.

Key words: Exploratory factor analysis, Spearman's Rho, Duration of Diabetes and HTN, Comorbidities, Social habits, Physical activity, Age.

INTRODUCTION

Drug Use Evaluation (DUE) also known as medication use evaluation is a technique that focuses on the usage of drugs and methods applied to attain the desired outcomes on patients based on a systematic evaluation of drug usage.[1] The drug usage pattern can explain the extent and complete profile of the drug usage, ongoing trends, quality of use audits and compare the drug usage with the national, regional and local guidelines and formularies. All aspects related to the drugs, their benefits and adverse effects can be studied via Drug use evaluation. Drug utilization data can be used for planning the drug production and expenditure. DUE can evaluate the extent of inappropriate prescribing of drugs. DUE aids help to the healthcare system for understanding, improving and interpreting prescription administration and use of medication.[2]

Diabetes Mellitus is one of the reasons for the increased morbidity and mortality rates and hypertension is the most commonly seen comorbidity in diabetic patients.[3] According to WHO hypertension is defined as a condition that causes a rise in the blood pressure within the blood vessels. According to the IDF Diabetes Atlas, 8th edition, 2017, there are 425 million people in this world who are enduring diabetes. There will be around 629 million people in the world who will be bearing with diabetes in 2045.[4]

The shift in the age distribution globally and amongst diabetic people and the number of deaths increasing due to diabetes, there is a need to understand the relationship between age and duration of diabetes.

Aims and Objectives

The purpose of this paper is to provide the socio-demographic characteristics of patients presented in the OPD and IPD medicine department. The aim of this study is to examine the relationship and conduct spearman correlation and factorial analysis of the duration of the disease with respect to socio-demographic characteristics of the patients.

MATERIALS AND METHODS

A prospective observational study was carried out in 250 patients who were suffering from both diabetes and hypertension. The study was conducted for a duration of 6 months from June 2018- November 2018. Approval for the study was given by the Institutional Ethics Committee (Approval No.: ECR/296/Indt/PB/2018/ISFCP/114) of I.S.F College of Pharmacy, Moga.

Participants who willingly involved in the study were asked to sign to consent form. The confidence interval of the study was selected as 95% with cut off interval (p=0.05) as a significant level. The sample size for the study was calculated using software Epiinfo (Stat Cal).[5] On the basis of the designed inclusion and exclusion criteria patients were enrolled for the study. Patients who suffered from both DM and HTN, patients (in-patients and out-patients) of both genders and patients who were willing to give consent for the study as per biomedical ethics were included for the study. Whereas, patients who were admitted in I.C.U, critical care units and other non-selected departments were not included in the study. The analysis was made using SPSS version 24. The results were evaluated based on the descriptive and analytical analysis. The descriptive study gives an outline of the basic characteristics of the data and analytical study was carried out to find the effect of the independent variable on the dependent variable. Factorial analysis and Spearman Correlation analysis was done in order to find out the relationship and correlation between variables as well as predicting the procedure for analysis to be made on a larger population.[6-9]

RESULTS

Descriptive analysis was directed to describe the demographic profile of the patients. During the study,

it was perceived that the mean age of the patients who suffered from both DM and HTN was found to be 58.23±12.04. Highest number of patients were found in the age group between 60-69 years that is 72 (31.2%) patients followed by 63 (25.2%) patients in the age group between 50-59 years and only 3 (1.2%) patients were falling under age group 20-29 years. The age group of the study was 22 to 86 years. Test for normality was conducted to check the normality distribution of the age of the patient.[10-12] As per Kolmogorov Smirnov (KS) and Shapiro-Wilk (SW) it was p=0.009[a], p=0.058[b] respectively. Age was found to be normally distributed because the p-value was found above the cut-off point with Skewness -0.200, Kurtosis 0.178 and standard error 0.762. A detailed description is given in Table 1.

The data was normally distributed with respect to the age of the patient as represented by histogram and Q-Q plot. A well-defined bell-shaped curve and straight line is observed and data points are near to the diagonal line in the Q-Q plot. Hence, mean (SD) was used to narrate the measurement of central tendency.[12] Figure 1 and Figure 2 represents histogram and normal O-Q plot of age distribution.

120 (48%) patients were male whereas 130 (52%) patients were female as observed in the study. Test for normality (Table 2) was conducted to check the normality distribution of age and gender of the patient. It was found that as per Kolmogorov Smirnov (KS) and Shapiro-Wilk (SW) was p=0.078[a], p=0.243[b] respectively for males and p=0.200[a], p=0.038[b] respectively for females.

Table 1: Tests of Normality.

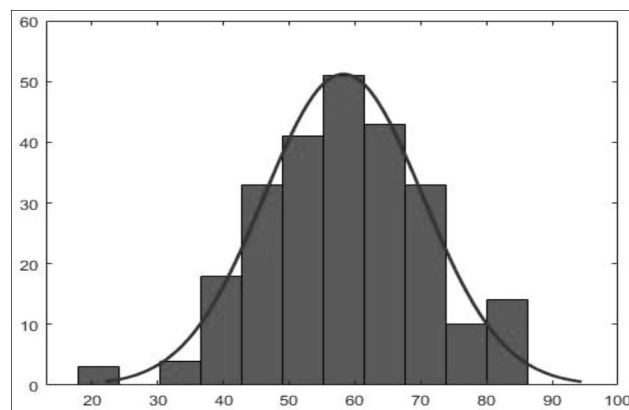| Variable	Kolmogorov-Smirnov[a]			Shapiro-Wilks		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Age of the patients	.066	250	.009	.989	250	.058

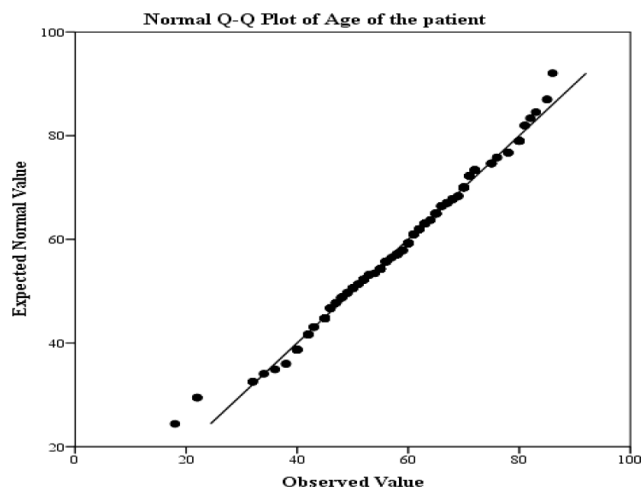**Figure 1: Histogram with normality plot of the age of the patient.**

Figure 2: Normal Q-Q plot of age distribution.

Table 2: Tests for Normality.

Parameter	Gender	Kolmogorov-Smirnova			Shapiro-Wilks		
		Statistic	Df	Sig.	Statistic	Df	Sig.
Age of the patient	Male	.077	120	.078	.986	120	.243
	Female	.063	130	.200*	.979	130	.038

The weight range of the study group was observed to be 42 to 89 kgs. A maximum number of patients that are 94 (37.6 %) were observed in the weight category of 70-79 kgs followed by 63 (25.2%) who were observed in the weight category of 60-69 kgs. Median weight of the patients who suffered from DM and HTN was found to be median (IQR) 70.5 (17.5). It was observed that the majority of the patients that is 50.8% were non-alcoholic and non-smokers. Whereas 33.6% of patients were alcoholics, 11.6 % of patients were smokers and 4.0% of patients were smokers as well as alcoholic. The highest number of patients who suffered from DM were newly diagnosed that is 60.4% patients. Whereas, the minimal number of 3.2 % of patients were observed to have DM for the past 5 years. Detailed representation is given in Table 3.

The highest number of patients who suffered from HTN were newly diagnosed that is 76.8 % of patients. Whereas, minimal number of 2.4 % of patients suffered from HTN from >10 years. A detailed description is given in Table 4.

Systolic BP

The highest number of patients were observed to be in the category of 140-159 mm Hg that is 63.2 % patients, followed by 20.4 % of patients who fall in the category of 120-139 mm Hg. Only 0.8 % of patients fall in the <120 mm Hg category. The minimum systolic blood pressure observed was 110 mm Hg whereas the maximum systolic blood pressure observed was 210 mm Hg.

Table 3: Distribution on the basis of Duration of diabetes mellitus.

Duration of DM	Frequency	Percentage
0-1 month (newly diagnosed)	151	60.4
< 5 years	8	3.2
5-10 years	46	18.4
>10 years	45	18.0
Total	**250**	**100.0**

Table 4: Distribution on the basis of Duration of hypertension.

Duration of hypertension	Frequency	Percent
0-1 month (newly diagnosed)	192	76.8
< 5 years	29	11.6
5-10 years	23	9.2
>10 years	6	2.4
Total	**250**	**100.0**

Diastolic BP

The highest number of patients were observed to be in the category of 90-99 mm Hg which is 58 % of patients followed by 80-89 mmHg that is 22.8 % patients. Very few number of patients were found to be in the category of <80 mm Hg that is 5.6 % patients.

Out of the total 250 patients who suffered from DM and HTN, it was noticed that 81.2 % of patients performed physical activity. While 18.8% of patients did not perform any physical activity. Maximum of 23.6 % were observed suffering from deficiency of Vitamin D followed by 21.2 % of patients who endured with GIT disorders.

Spearman correlation coefficient test

A Spearman correlation coefficient also is known as spearman rank correlation or Spearman's Rho is a non parametric test that is used to determine the strength as well as the direction of the association between two numeric variables. It defines whether there is a monotonic relation between two continuous or ordinal variables.[13] A positive correlation coefficient denotes a positive relationship between two variables i.e when the value of one of the variables increases, directly the value of another variable also increases whereas a negative correlation coefficient denotes a negative relationship. If a correlation coefficient comes out to be zero it indicates that there is no relation between the two variables, although a nonlinear relation might exist. If p value= <.0.05 Represents positive correlation and p value= >.0.05 represents negative correlation.[14] We have studied spearman correlation for checking the effect of age on the duration of HTN. Table 5 gives a detailed

description of the correlation coefficient between the age and duration of HTN.

It was observed that there is a significant moderate positive correlation between age and duration of HTN ($r = 0.408$, $P = 0.001$). Figure 3 represents a scatter plot of the age of the patient and the duration of HTN and shows a moderate positive correlation.

Factorial analysis

Factor analysis is a statistical method used to determine the relationship between components of a dataset. It is performed to study one set of variables and refer to the correlation within this set. Thereby, the correlation structure of various several other sets of variables can be studied. Here the variance of a factor is denoted by an eigenvalue that is enumerated as the sum of its squared factor loadings for all the items.[15,16] This method helps to extract maximum common variance from all variables and assigns them into a common score and we can use this score for further analysis. Here, in this paper exploratory factor analysis has been done which assumes that any indicator or variable within the data set

may be associated with any of the factors. Univariate and multivariate normality within the data set was confirmed before intiating factor analysis.

The correlation matrix represents an abridged version of R-matrix. The top half of the table restrains Pearson's correlation coefficient between all pairs of question whereas the bottom half contains the one-tailed significance of this coefficient. The value of significance was scanned.[17,18] Table 6 denotes both correlation and significance.

Value of Kaiser-Meyer-Olkin close to 1 indicates the pattern of correlation and is compact therefore, factor analysis should yield well founded facts. Here the value of KMO is 0.51 which indicates mediocre value. Bartlett's test of sphericity measure was used to test the null hypothesis which was found to be significant because of *p=0.001*.

The principle behind table of communalities before and after extraction is based on the assumption that all variance is common, therefore before the extraction, the communalities are all 1. Here, Table 7 shows that 67.9% of the variance linked with parameter 1 is common, by 90.1% of the variance linked with parameter 2 is common.

Correlation matrix

Refer to Figure 4 and Table 8 for the correlation matrix and scree plot. The scree plot specifies the point of inflextion on the curve. The curve begins to tail off after 1 factor but there is another drop after 2 factors before a stable plateau is reached.

Factor Analysis

The correlation matrix table represents Pearson's correlation coefficient between all pairs of the question in the top half of the table whereas the bottom half denotes the one-tailed significance of this coefficient. Here the value of KMO is 0.49 (Table 9) that indicates acceptable value. Bartlett's test of sphericity test the null hypothesis is significant because of *p=0.001*.[19,20]

Table 10 shows both the correlation and significance and

Table 5: Spearman Correlation Coefficient.

Parameter	Test	Age of the patient	Duration of hypertension
Age of the patient	Correlation Coefficient	1.000	.408
	Sig. (2-tailed)	.	.001
	N	250	250
Duration of hypertension	Correlation Coefficient	.408	1.000
	Sig. (2-tailed)	.001	.
	N	250	250

Figure 3: Scatter dot plot of Age of the patient and Duration of HTN.

Table 6: KMO and Bartlett's Test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.518
Bartlett's Test of Sphericity	Approx. Chi-Square	374.802
	Df	21
	Sig.	.001.

Table 7: Communalities.

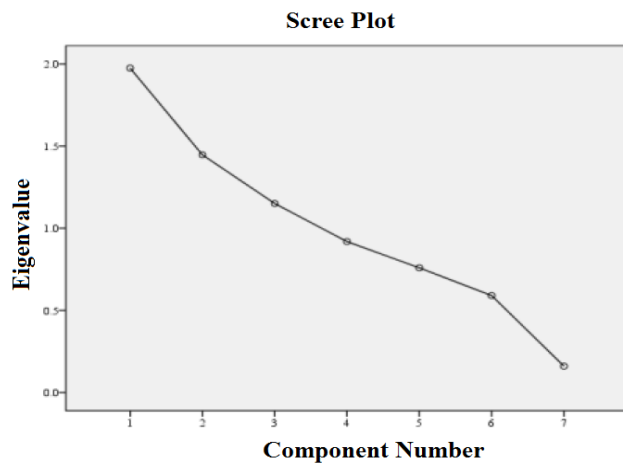| Parameters	Initial	Extraction
Age of the patient	1.000	.679
Gender of the patient	1.000	.901
Weight of the patient	1.000	.470
Duration of diabetes mellitus	1.000	.583
Duration of hypertension	1.000	.494
Social habits of the patient	1.000	.911
Physical activity performed	1.000	.534

Scree Plot

Figure 4: Scree plot shows that the curve begins to tail off after 1 factor but there is another drop after 2 factors before a stable plateau is reached.

Figure 5 represents a scree plot. The scree plot indicates the point of inflextion on the curve. The curve begins to tail off after 1 factor but there is just another drop after 2 factors before a stable plateau is reached.

Here, Table 11 shows that 64.9% of the variance associated with parameter 1 is commonly followed by 56.4% of the variance associated with parameter 2 is common. Detailed representation is given in the table below.

DISCUSSION

It was observed that maximum number of diabetic hypertensive patients fall between the age group of 60-69 years and mean(SD) age of the patients was observed to be 58.23±12.04. Bela Patel *et al.* (2013) concluded that maximum patients were found to be in the age group of 35-78 years and mean(SD) age of the patients was observed to be 56.8 ± 10.5 years.[21] Our results indicated that diabetes is more prevalent in females (52%) than in males (48%). Whereas, J Shah *et al.* (2012) concluded that

## Table 9: KMO and Barlette's Test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.496
Bartlett's Test of Sphericity	Approx. Chi-Square	49.080
	Df	15
	Sig.	.001

Table 8: Correlation Matrix.

	Parameters	Age	Gender	Weight	Duration of DM	Duration of HTN	Social habits	Physical activity performed
Correlation	**Age**	1.000	-.077	.291	-.097	.160	.163	-.278
	Gender	-.077	1.000	-.037	-.045	.042	-.834	.112
	Weight	.291	-.037	1.000	.006	.015	.056	-.102
	Duration of DM	-.097	-.045	.006	1.000	.098	.069	-.137
	Duration of HTN	.160	.042	.015	.098	1.000	-.041	-.199
	Social habbits	.163	-.834	.056	.069	-.041	1.000	-.159
	Physical activity performed	-.278	.112	-.102	-.137	-.199	-.159	1.000
Sig. (1-tailed)	**Age**		.114	.000	.064	.006	.005	.000
	Gender	.114		.282	.242	.253	.000	.039
	Weight	.000	.282		.461	.410	.191	.054
	Duration of DM	.064	.242	.461		.061	.140	.015
	Duration of HTN	.006	.253	.410	.061		.262	.001
	Social habbits	.005	.000	.191	.140	.262		.006
	Physical activity performed	.000	.039	.054	.015	.001	.006	

Table 10: Correlation Matrix.

	Parameter	Age	Weight	Systolic BP	Diastolic BP	Social habits	Comorbidities
Correlation	Age	1.000	.291	.093	-.039	.163	-.002
	Weight	.291	1.000	.095	.025	.056	.160
	Systolic blood pressure	.093	.095	1.000	.165	-.030	.053
	Diastolic blood pressure	-.039	.025	.165	1.000	.006	-.042
	Social habbits	.163	.056	-.030	.006	1.000	.055
	Comorbidities	-.002	.160	.053	-.042	.055	1.000
Sig. (1-tailed)	Age		.000	.072	.272	.005	.491
	Weight	.000		.067	.348	.191	.006
	Systolic blood pressure	.072	.067		.004	.321	.203
	Diastolic blood pressure	.272	.348	.004		.464	.253
	Social habbits	.005	.191	.321	.464		.195
	Comorbidities	.043	.006	.203	.253	.195	
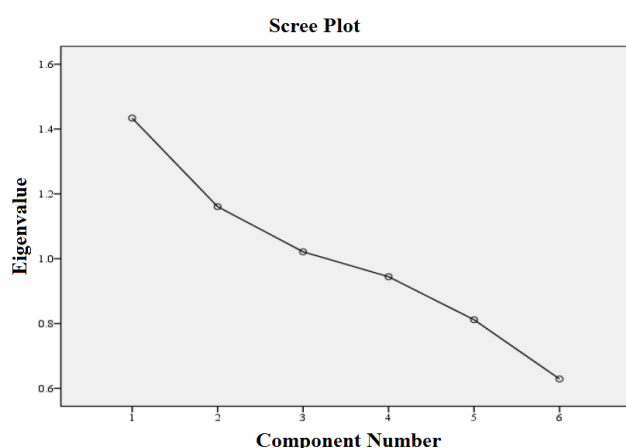
Scree Plot

Figure 5: Scree plot indicates the point of inflexion on the curve.

Table 11: Communalities.

Parameter	Initial	Extraction
Age of the patient	1.000	.649
Weight of the patient	1.000	.564
Systolic blood pressure	1.000	.589
Diastolic blood pressure of the patient	1.000	.613
Social habits of the patient	1.000	.431
Comorbidities	1.000	.770

males were more prevalent to suffer from diabetes than females.[22] In this study, the median (IQ) of the weight was observed to be 70.5 (17.5) whereas, Neil *et al.* (2015) concluded median weight of his study to be 80.5.[23] In the present study the mean duration of being diagnosed with diabetes was 2.69±4.46 years compared to 6.6±5.8 years in alike study performed by Mandana Moradi *et al.* (2016).[24] The median(IQ) duration of HTN in our study is 2(2) whereas Dhanraj *et al.* (2012) reported a median duration to be 4 (1-10) years.[25] Tingling in extremities

and weakness was the most common symptom observed during diagnosis in 29 (11.6%) patients whereas Bela Patel *et al.* (2013) reported weakness and fatigue in 71.2% of patients.[26] Further analysis was done to study the effect of variables on each other.

CONCLUSION

In this study, it was observed that there is a significant moderate positive correlation between age and duration of HTN. Multiple factors are observed to be directly associated with the socio-demographic factors of the patients. During the study, it was observed that age, social habits and physical activity also has a wide impact on the duration of the disease itself. The analysis made in this study can be useful for the purpose of clinical guideline development and profit the policy makers and health care professionals of India to earmark the resources more systematically to address the needs and care for diabetic patients.

Limitations Of Our Study

The present study did not include all the diabetic hypertensive patients who visited the hospital during the study period because of logistical issues. Their enrollment would have given many clear data without exclusion of any patient.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ABBREVIATIONS

DM: Diabetes Mellitus; **HTN:** Hypertension; **DUE:** Drug utilization evaluation; **BP:** Blood pressure.

SUMMARY

- Moderate positive correlation between age and duration of HTN.

- Age, social habits and physical activity also has a wide impact on the duration of the disease.

- Vitamin D deficiency was the major comorbidity observed followed by GIT disorders.

REFERENCES

1. Moradi M, Sarah M. Drug use evaluation of diabetes mellitus in non-hospitalized patients. International Journal of Pharmacy and Pharmaceutical Sciences. 2019;8(8):337-41.

2. Shalini S, Ravichandran V, Mohanty BK, Dhanaraj SK, Saraswathi R. Drug utilization studies: An overview. Inter J Pharmaceut Sci Nanotechnol. 2010;31:803-10.

3. Alavi NM, Alami L, Taefi S, Gharabagh GS. Factor analysis of self-treatment in diabetes mellitus: A cross-sectional study. BMC Public Health. 2011;11(1):761.

4. Monika K, Harbans LK, Jagminder KB. Drug Utilization Evaluation in Diabetic and Hypertensive Patients and Adherence to Guidelines at a Teaching Hospital. EC Endocrinology and Metabolic Research. 2019;4(9):45-54.

5. Sharma A, Joshi N, Baldi A, Sharma DK. Socio-Demographic Characteristics and Drug Related Problems of Patients Presenting to the Emergency Department: General Linear Model and Factorial Analysis. J Pharma Care Health Sys. 2018;5(194):2376-0419.

6. Schutte AE, Rooyen JMV, Huisman HW, Kruger HS, DeRidder JH. Factor analysis of possible risks for hypertension in a black South African population. Journal of Human Hypertension. 2003;17(5):339.

7. Ayubi E, Khalili D, Delpisheh A, Hadaegh F, Azizi F. Factor analysis of metabolic syndrome components and predicting type 2 diabetes: Results of 10-year follow-up in a Middle Eastern population. J Diabetes. 2015;7(6):830-8.

8. Martin CE, Renn BN, Winderman KE, Hundt N, Petersen NJ, Naik AD, *et al*. Classifying diabetes-burden: A factor analysis of the Problem Areas in Diabetes Scale. Journal of Health Psychology. 2018;23(6):882-8.

9. Alavi NM, Alami L, Taefi S, Gharabagh GS. Factor analysis of self-treatment in diabetes mellitus: A cross-sectional study. BMC Public Health. 2011;11(1):761.

10. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: A guide for non statisticians. International Journal of Endocrinology and Metabolism. 2012;10(2):486.

11. Laerd Statistics. Testing for normality using SPSS. 2013;1-5.

12. Rahman J. Brief guidelines for methods and statistics in medical research. Springer Singapore. 2015.

13. https://www.statisticssolutions.com/spearman-rank-correlation

14. 14.https://statistics.laerd.com/spss-tutorials/spearmans-rank-order-correlation-using-spss-statistics.php

15. Schneeweiss H, Mathes H. Factor analysis and principal components. Journal of Multivariate Analysis. 1995;55(1):105-24.

16. https://www.statisticssolutions.com/factor-analysis-sem-factor-analysis

17. Yong AG, Pearce S. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in Quantitative Methods for Psychology. 2013;9(2):79-94.

18. Williams JS, Child D. The Essentials of Factor Analysis. Cassell Educational. 2003;8.

19. Oh JH, Thor M, Olsson C, Skokic V, Jörnsten R, Alsadius D, *et al*. A factor analysis approach for clustering patient reported outcomes. Methods of Information in Medicine. 2016;55(05):431-9.

20. Alavi NM, Alami L, Taefi S, Gharabagh GS. Factor analysis of self-treatment in diabetes mellitus: A cross-sectional study. BMC Public Health. 2011;11(1):761.

21. Patel B, Oza B, Patel KP, Malhotra SD, Patel VJ. Pattern of antidiabetic drugs use in type-2 diabetic patients in a medicine outpatient clinic of a tertiary care teaching hospital. Int J Basic Clin Pharmacol. 2013;2(4):485-91.

22. Shah J, Khakhkhar T, Bhirud S, Shah RB, Date S. Study of utilization pattern of anti-hypertensive drugs in hypertensive diabetic patients with or without reduced renal function at tertiary care teaching hospital. Int J Med Sci Public Health. 2013;2(2):175-80.

23. DeFine ON, Siersma VD, Køster-Rasmussen R, Heitmann BL, Waldorff FB. Weight changes following the diagnosis of type 2 diabetes: The impact of recent and past weight history before diagnosis. results from the Danish Diabetes Care in General Practice (DCGP) study. PLoS One. 2015;10(4):e0122219.

24. Moradi M, Sarah M. Drug use evaluation of diabetes mellitus in non-hospitalized patients. International Journal of Pharmacy and Pharmaceutical Sciences. 2019;8(8):337-41.

25. Dhanaraj E, Raval A, Yadav R, Bhansali A, Tiwari P. Prescription pattern of antihypertensive agents in T2DM patients visiting tertiary care centre in North India. International Journal of Hypertension. 2012.

26. Patel B, Oza B, Patel KP, Malhotra SD, Patel VJ. Pattern of antidiabetic drugs use in type-2 diabetic patients in a medicine outpatient clinic of a tertiary care teaching hospital. Int J Basic Clin Pharmacol. 2013;2(4):485-91.