Evaluating the Effects of Treatment Type and Age Group on Blood Glucose Levels in Patients with Diabetes: A Factorial Experimental Design Study.

Hakeem – Adekunle[1], *Oladimeji – Adewuyi*[2], *Kayode – Okunola*[3]

December 2024

1 State Of The Problem

Diabetes management remains a global health challenge (WHO, 2023), requiring personalized treatment strategies to optimize patient outcomes. Blood glucose control, a key indicator of diabetes management, is influenced by various factors, including treatment methods and demographic characteristics such as age. However, the interaction between these factors is not well understood, making it difficult to determine the most effective treatment for specific patient groups.

This study aims to evaluate the main effects of Treatment Type (oral medication, insulin therapy, and combined therapy) and Patient Age Group (18–30 years, 31–50 years, and 51–70 years) on blood glucose levels. Additionally, the study investigates whether the effectiveness of different treatment types varies across age groups. Understanding these interactions can provide valuable insights into tailoring treatment plans to improve diabetes outcomes.

By employing a factorial experimental design and analyzing the results using 3^2 factorial design, this study seeks to address the following research questions:

- Does treatment type significantly affect blood glucose levels?
- Does age group significantly affect blood glucose levels?
- Is there a significant interaction between treatment type and age group in influencing blood glucose levels?

2 Methodology

2.1 Data Collection Process

The data for this study were collected to evaluate the effects of treatment type and patient age on blood glucose levels. Participants were randomly selected and assigned to one of three treatment types, and their age was recorded to classify them into three age groups. Following the administration of the assigned treatment, blood glucose levels were measured using standardized medical instruments to ensure accuracy and reliability. Each treatment type and age group combination included five replicates, resulting in 45 observations. Each observation recorded the treatment type, the participant's age group, and their post-treatment blood glucose level. This factorial design allowed for the analysis of the main effects of treatment type and age, as well as the interaction between these two factors, in influencing blood glucose levels.

Table 1: Data

Obs	Treatment_type	Patients_age	glucose_level
1	1	1	309
2	1	1	290
3	1	1	285
4	1	1	300
5	1	1	310
6	1	2	295
7	1	2	290
8	1	2	295
9	1	.	1
10	1	3	320
11	1	3	219
12	1	3	316
13	1	3	299
14	1	3	371
15	2	1	290
16	2	1	230
17	2	1	235
18	2	1	240
19	2	1	260
20	2	2	240
21	2	2	235
22	2	2	220
23	2	2	225
24	2	2	321
25	2	3	200
26	2	3	333
27	2	3	301
28	2	3	296
29	2	3	293
30	3	1	297
31	3	1	317
32	3	1	300
33	3	1	311
34	3	1	316
35	3	2	322
36	3	2	311
37	3	2	319
38	3	2	321
39	3	2	302
40	3	3	297
41	3	3	296
42	3	3	314
43	3	3	295
44	3	3	304

2.1.1 3^2 Factorial Design

In this research, we introduced a 3^2 factorial experimental design, meaning that we include two factors, each at three levels. This design allows for the investigation of the main effects of the two factors as well as their interaction effect.

This factorial design enables the study of how each factor independently affects the response variable (e.g., blood glucose level) and whether there is an interaction between the two factors.
Mathematically;
In a 3^2 factorial design, the response variable Y_{ijk} can be modeled as:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Where:

- μ: Overall mean.
- α_i: Effect of the i-th level of factor A (Treatment type).
- β_j: Effect of the j-th level of factor B (Patient age group).
- $(\alpha\beta)_{ij}$: Interaction effect between the i-th level of factor A and the j-th level of factor B.
- ϵ_{ijk}: Random error term, assumed to be normally distributed with mean 0 and variance σ^2.

2.1.2 Hypotheses

Main Effects:

1. Effect of Treatment Type (Factor A):

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

$$H_a : \text{At least one } \alpha_i \neq 0$$

2. Effect of Age Group (Factor B):

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_a : \text{At least one } \beta_j \neq 0$$

Interaction Effect:

Interaction between Treatment Type and Age Group:

$$H_0 : (\alpha\beta)_{ij} = 0 \text{ for all } i \text{ and } j$$

$$H_a : (\alpha\beta)_{ij} \neq 0 \text{ for at least one } i, j$$

Typically, $\alpha = 0.05$ is used, indicating a 5% chance of rejecting the null hypothesis when it is true. The critical F-value is determined based on the degrees of freedom for the numerator (factor levels or interaction levels) and denominator (error term) from the F-distribution table at the chosen significance level.

F-Ratios

- For Treatment Type (Factor A):

$$F_A = \frac{\text{MS}_A}{\text{MS}_E}$$

$\text{MS}_A = \frac{\text{SS}_A}{df_A}$, $\text{MS}_E = \frac{\text{SS}_E}{df_E}$

- For Age Group (Factor B):

$$F_B = \frac{\text{MS}_B}{\text{MS}_E}$$

- For Interaction:

$$F_{AB} = \frac{\text{MS}_{AB}}{textMS_E}$$

Where SS_A, SS_B, SS_{AB}, and SS_E are the sums of squares for the respective components.

3 Results

3.0.1 Model Summary

The results in the table below returned a P-value of 0.0111 which is less than the chosen significant level (0.05), meaning that the overall model is statistically significant (F(8, 34) = 3.03, p = 0.0111), and the model is a good fit. (Table 1).

3.0.2 Effects Estimation

In Table 2, we can see that, only the factor; treatment type has a p-value less than the chosen significance level (0.05) while the patient's and the interaction effect has a p-value greater than the significance level, we can say that, the factor; treatment type is significant and no interaction effect exists between the factors.

3.1 Model Validation: Normality of the Residual

From figure 3, it can be seen that the points are scatters in both side of the straight lines, implying that, some residuals are not close to normal.

Table 1: Model Summary

Dependent Variable: glucose_level

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	23401.90078	2925.23760	3.03	0.0111
Error	34	32841.86667	965.93725		
Corrected Total	42	56243.76744			

Table 2: ANOVA Results

R-Square	Coeff Var	Root MSE	glucose_level Mean
0.416080	10.74120	31.07953	289.3488

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment_type	2	18463.51103	9231.75552	9.56	0.0005
Patients_age	2	1213.33451	606.66725	0.63	0.5397
Treatment*Patients_a	4	3725.05524	931.26381	0.96	0.4398

3.2 Post Hoc Test

Since the treatment type is the only significant factor, the purpose of a post hoc test is to identify which specific levels of a factor differ significantly from each other.

Figure 1: Interaction Plots

Figure 2: Distribution of the Treatment Type

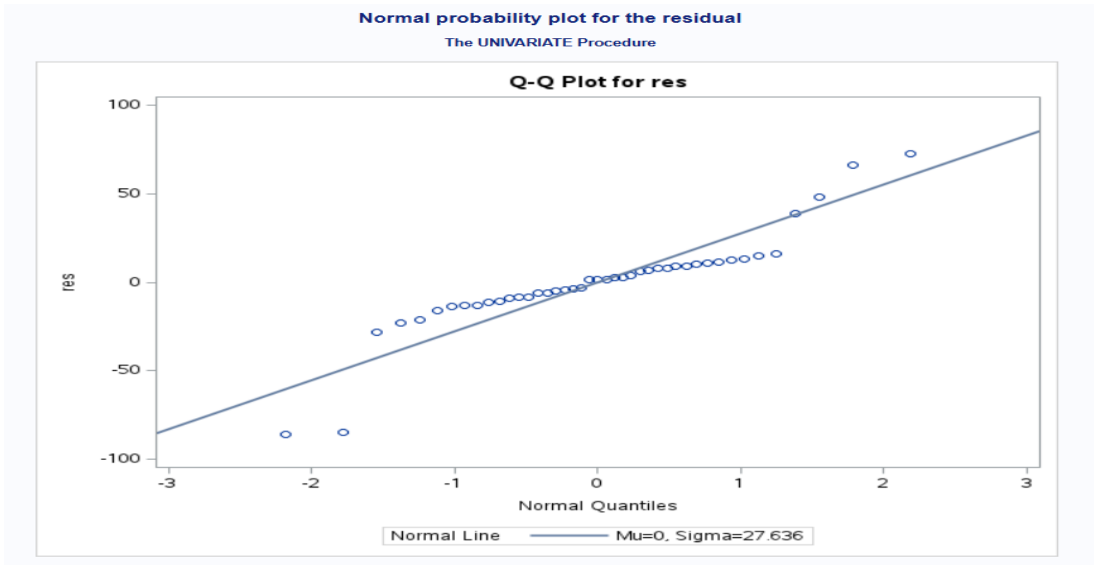

Figure 3: QQ Plot of the Residual Terms

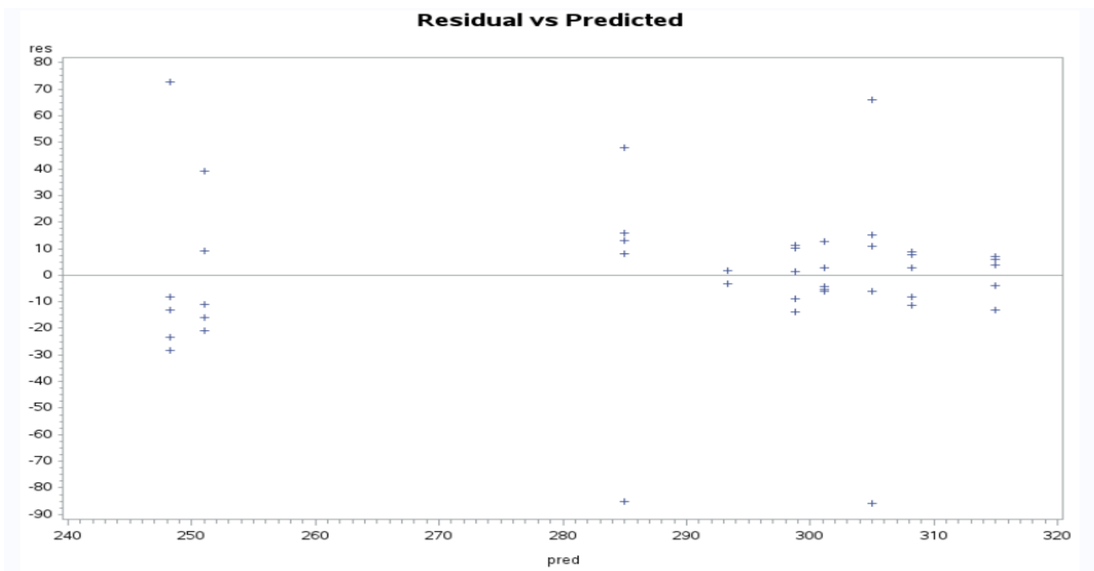

Figure 4: Residual Vs Predicted Plot

3.2.1 Pairwise Comparisons

Treatment 2 appears to have significantly different effects on glucose levels compared to Treatments 1 and 3. However, there is no significant difference between Treatments 1 and 3. This analysis provides clarity on the relative effectiveness of each treatment type. (Table 3).

Table 3: Result of the Pairwise Comparison Test

Tukey's Studentized Range (HSD) Test for glucose_level

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	34
Error Mean Square	965.9373
Critical Value of Studentized Range	3.46544

Comparisons significant at the 0.05 level are indicated by ***.				
Treatment_type Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
3 - 1	8.21	-20.65	37.07	
3 - 2	46.73	18.92	74.54	***
1 - 3	-8.21	-37.07	20.65	
1 - 2	38.52	9.66	67.38	***
2 - 3	-46.73	-74.54	-18.92	***
2 - 1	-38.52	-67.38	-9.66	***

4 Conclusion

The results indicate that the study successfully examined the impact of different treatments on glucose levels, achieving its objective of determining whether significant differences exist between treatment groups. The ANOVA analysis confirms that the treatment groups influence glucose levels, as the model is statistically significant. Post-hoc comparisons using Tukey's HSD further reveal specific significant differences between treatment groups, highlighting notable variations in their effects on glucose levels. These findings provide evidence that the treatments have a measurable and meaningful impact, supporting the need for further investigation into the underlying mechanisms and potential practical applications of these results.

5 Appendix

5.0.1 SAS Scripts

```
data tube;
input Treatment_type Patients_age glucose_level;
datalines;
1 1 309
1 1 290
1 1 285
1 1 300
1 1 310
1 2 295
1 2 290
1 2 295
1 2 308
1 2 303
1 3 320
1 3 219
1 3 316
1 3 299
1 3 371
2 1 290
2 1 230
2 1 235
2 1 240
2 1 260
2 2 240
2 2 235
2 2 220
2 2 225
2 2 321
2 3 200
2 3 333
2 3 301
2 3 298
2 3 293
3 1 297
3 1 317
```

```
3 3 296
3 3 314
3 3 295
3 3 304
;
run;
proc glm data=tube;
class  Treatment_type Patients_age;
model glucose_level = Treatment_type Patients_age Treatment_type*Patients_age;
means Treatment_type / tukey; /* Tukey's Test */
output out=values r=res p=pred;
run;
Title1 "Normal probability plot for the residual";
proc univariate data=values normal ;
var res;
qqplot res/normal(L=1 mu=0 sigma=est);
run;
Title "Residual vs Glass";
proc gplot data=values;
plot res*Treatment_type/haxis=axis1 vref=0;
run;

Title "Residual vs Phosphor";
proc gplot data=values;
plot res*Patients_age/haxis=axis1 vref=0;
run;
Title "Residual vs Predicted";
proc gplot data=values;
plot res*pred/haxis=axis1 vref=0;
run;
```