AUCTORES
Globalize your Research

Journal of Endocrinology and Disorders

Rabindra Nath Das *

Open Access

Research Aticle

Effects of Body Mass Index on the Polycystic Ovary Syndrome Women

Mahashweta Das [1], Prabir Chakraborty [2], Rabindra Nath Das [3]*

[1]Department of History, The University of Burdwan, Burdwan, West Bengal, India.

[2]Department of Political Science, Govt. Degree College Gandacherra, Dhalai, Tripura, India.

[3]Department of Statistics, The University of Burdwan, Burdwan, West Bengal, India.

***Corresponding Author:** Rabindra Nath Das, 3Department of Statistics, The University of Burdwan, Burdwan, West Bengal, India.

Received Date: 10 October 2025 | **Accepted Date:** 18 October 2025 | **Published Date:** 29 October 2025

Citation: Mahashweta Das, Prabir Chakraborty, Rabindra N. Das, (2025), Effects of Body Mass Index on the Polycystic Ovary Syndrome Women, *J. Endocrinology and Disorders*, 9(5): **DOI:10.31579/2640-1045/225**

Abstract

Polycystic ovary syndrome (PCOS) is a usual endocrine medical situation that assails a large number of women in adolescence and reproductive age. The PCOS distribution over different body mass index (BMI) grades can vary, and the current research has shown the effects of BMI on PCOS women, using 1000 real observations, and the data set is in the site: https://www.kaggle.com/datasets/samikshadalvi/pcos-diagnosis-dataset The BMI values analysis findings are developed herein applying statistical joint generalized linear models (JGLMs). It is developed herein that mean BMI is negatively associated with the joint interaction effect (JIE) of the subject's testosterone (TET) levels and menstrual irregularity (MIT) i.e., TET*MIT (P<0.0001), while it is positively associated with both TET (P=0.0025) and MIT (P<0.0001). Mean BMI is negatively associated with the JIE of the subject's antral follicle count (AFC) values and MIT i.e., AFC*MIT (P<0.0001), while it is positively associated with both AFC (P=0.0387) and MIT (P<0.0001). Mean BMI is negatively associated with the JIE of the subject's TET levels and AFC values i.e., TET*AFC (P=0.0253), while it is positively associated with both TET (P=0.0025) and AFC (P=0.0387). Mean BIM value is positively associated with the JIE of TET levels and the subject's PCOS diagnostic status i.e., TET*PCOS (P<0.0001), while it is positively associated with TET (P=0.0025) and negatively with PCOS (P=0.0661). Mean BIM value is positively associated with the JIE of AFC values and the subject's PCOS diagnostic status i.e., AFC*PCOS (P<0.0001), while it is positively associated with AFC (P=0.0387) and negatively with PCOS (P=0.0661). BMI values' variance is negatively associated with age (P=0.0839), PCOS (P<0.0001) and AFC (P=0.1170). The article has shown that BMI value has different significant JIEs on PCOS women. The current outcomes regarding the BMI values may be instrumental for the PCOS women, practitioners and researchers. It concludes that BMI along with MIT, TET and AFC has multiple effects on PCOS women.

Key Words: body mass index; Antral follicle count; Testosterone levels; Joint mean-variance model; Polycystic ovary syndrome

Introduction

Polycystic ovary syndrome (PCOS) is a usual endocrine medical situation that assails a large number of women in adolescence and reproductive age. Many articles [1,2] have pointed out that approximately 8-16% of women are affected by PCOS globally in their reproductive age group. The PCOS distribution over different body mass index (BMI) grades can vary, many research articles have tried to focus on the association between BMI and PCOS women [3-5].

PCOS can have very complex interaction effects with various factors such as BMI, testosterone (TET) levels, menstrual irregularity (MIT), antral follicle count (AFC), age including lifestyle choices [6,7]. Obesity is highly prevalent in PCOS women. There is a preferential androgenic distribution pattern of body fat in PCOS women [8]. Reproductive functional problems such as MIT and infertility are more prevalent in obese women [8-10]. Obesity is also directly associated with PCOS, which assails 6–12% of women of their reproductive age [11].

The association between elevated BMI and individual phenotypic characteristics of the Rotterdam criteria that discriminate against PCOS remains hazy [12]. For instance, a meta-analysis among PCOS women shows that hirsutism, as quantified by modified Ferriman-Gallwey score, was elevated only when comparing obese women versus overweight, but not when comparing obese women versus normal weight [13]. Roles of BMI on features such as MIT, and AFC remain unclear, especially in

healthy women. From many articles [7,9,10,11, 14], it is suspected that BMI affects these above-mentioned individual phenotypic features (MIT, AFC, TET, PCOS diagnosis status) of women with PCOS or without PCOS.

In previous research articles [8,12,13], the effects of BMI on PCOS women are unclear. Presently, some advanced research tools such as machine learning, statistical modelling, data mining etc. are employed in PCOS data analysis [10,12,15]. Several machine learning algorithms such as Random Forest, locally weighted learning, Decision table, Multilayer perceptron, Random tree, etc. are employed in the PCOS data analysis [2,15,16]. Several common statistical techniques such as testing of hypotheses, simple correlation & regression, analysis of variance etc. are applied in PCOS data analysis that are not suitable for positive, non-normal and non-constant variance PCOS data sets [2,4,8,11].

The present BMI response in the considered PCOS data set is a non-constant variance dependent variable, which is positive and non-normal. Best of our knowledge, most of the previous manuscripts did not consider the response BMI in PCOS data sets as a heteroscedastic and non-normal random variable. Therefore, most of the BMI analysis reports of PCOS data sets invite several doubts and debates. The effects of BMI on PCOS women are little studied adopting advanced probabilistic modeling. The present BMI study manuscript for PCOS data set searches the following research statistical hypotheses that are connected with and without PCOS women.

- Does BMI associate with irregular menstrual cycles, age, TET levels, AFC values and polycystic ovarian morphology of PCOS women?
- If it is affirmative, how can one derive the most probable BMI grades association model?
- What are the most probable BMI grades statistical model?
- What are the effects of BMI grades on PCOS women?

The article studies the above BMI grades examination research hypotheses adopting the following paragraphs such as materials & methods, statistical analysis & results, discussions, and conclusions. Statistical mean & variance models of the response BMI are revealed in Table 1, based on the PCOS data set, which is marked in the materials section. Mean and variance joint statistical model of the response BMI is obtained using joint generalized linear models (JGLMs), which is shortly revealed in the methods section. Response BMI modelling outcomes are presented in the result section, while the BMI modelling outcomes are presented in the discussion section. The BMI analysis's main information is noted in the conclusions section.

2. Materials and Methods

2.1. Materials

The current BMI study dataset is a sample of 1000 women subjects with PCOS and without PCOS, while PCOS is a usual hormonal endocrine disorder assailing woman of their reproductive age. The considered PCOS sample data set contains six correlated characters which are primarily connected with PCOS diagnosis. These six characters are considered as the valued insights into the subjects' medical health situations, and they can be employed for exploratory data analysis such as statistical modelling, feature engineering and machine learning for identifying PCOS diagnosis status. The considered PCOS data is available in the site : https://www.kaggle.com/datasets/samikshadalvi/pcos-diagnosis-dataset

The under study PCOS data set comprises six features such as the sample unit woman's body mass index (BMI), age, testosterone level (TET), antral follicle count (AFC), menstrual irregularity (MIT) (0=No, 1= Yes) and polycystic ovary syndrome (PCOS) (0=No, 1=Yes) diagnosis status. The sample study women are taken in their reproductive age. In the present study, BMI is the response variable, which is a body fat measure that is computed based on height and weight, and it is commonly ranging from 18 to 35. It is computed using the weight (in Kg) and height (in meter), and it is defined as BMI= Weight (kg) / Height (m2). BMI is a widely used screening tool that provides a simple numeric measure of an individual's weight in relation to their height as stated above. It is generally used to categorize individuals as underweight, normal weight, overweight, or obese. BMI is an important risk factor for many diseases such as diabetes, cardiovascular disease, and hormonal disorders such as PCOS etc. The different categories of individuals based on BMI are as follows. Individuals are categorized as underweight class when BMI values < 18.5. If the BMI value lies between 18.5 to 24.9, the individuals are categorized as normal weight. If the BMI value lies between 25.0 to 29.9, the individuals are categorized as overweight. If the BMI value lies between 30.0 to 34.9, the individuals are categorized as obesity (class I). If the BMI value lies between 35.0 to 39.9, the individuals are categorized as obesity (class II). Individuals are categorized as extreme obesity (class III) when BMI values > 39.8. The five BMI's explanatory variables are age, MIT (0 = No, 1 = Yes), TET levels, AFC values and the subject's PCOS diagnosis status (0=No, 1=Yes).

2.2 Statistical Methods

The current study takes into account the BMI grades as the response random variable, and it is to be modeled with the remaining five variables such as age, TET levels, MIT, AFC values and PCOS diagnosis status. The response BMI is identified as a non-normally and non-constant variance distributed random variable. The BIM's variation can't be stabilized by any proper transformation, so BIM value is modeled herein using joint generalized linear models (JGLMs) considering both the Gamma and Log-normal distributions, which is well described in [17-20]. Joint mean & variance models i.e., JGLMs are well described in the book by Lee et al. [17] and in the book by Das [18]. A short note of JGLMs for BMI values under both the Log-normal and Gamma distribution is displayed as follows.

JGLMs for Log-normal distribution: For the positive response Y_i (=BMI) with $E(Y_i=BMI) = \mu_i$ (mean) and $Var(Y_i=BMI) = \sigma_i^2 \mu_i^2 = \sigma_i^2 V(\mu_i)$ say, where σ_i^2's are dispersion parameters and $V(\cdot)$ reveals the variance function. Generally, log transformation $Z_i = \log (Y_i=BMI)$ is adopted to stabilize the variance $Var(Z_i) \approx \sigma_i^2$, but the variance may not always be stabilized [21]. For developing a BMI improved model, JGLMs for the mean and dispersion are considered. For the response BMI, assuming log-normal distribution, JGL mean and dispersion models (with $Z_i = \log (Y_i=BMI)$) are as follows:

$E(Z_i)= \mu_{zi}$ and $Var(Z_i) = \sigma_{zi}^2$,

$\mu_{zi}=x_i^t \beta$ and $\log(\sigma_{zi}^2)= g_i^t \gamma$,

where x_i^t and g_i^t are the explanatory factors/variables vectors of BMI values associated with the mean regression coefficients β and dispersion regression coefficients γ, respectively.

JGLMs for Gamma distribution: In the above stated Y_i's (=BMI), the variance has two portions such as $V(\mu_i)$ (based on the mean parameters μ_i's) and σ_i^2 (free of μ_i's). The variance function $V(\cdot)$ displays the GLM family distributions. For instance, if $V(\mu)= 1$, it is normal, Poisson if $V($

$\mu) = \mu$, and gamma if $V(\mu) = \mu^2$ etc. Gamma JGLMs mean and dispersion models of GLU are as follows:

$$\eta_i = g(\mu_i) = x_i^t \beta \text{ and } \varepsilon_i = h(\sigma_i^2) = w_i^t \gamma,$$

where $g(\cdot)$ and $h(\cdot)$ are the GLM link functions attached with the mean and dispersion linear predictors respectively, and x_i^t, w_i^t are the explanatory factors/variables vectors of BMI values attached with the mean and dispersion parameters respectively. Maximum likelihood (ML) method is used for estimating the mean parameters, while the restricted ML (REML) method is applied for estimating the dispersion parameters, which are explicitly stated in the book by Lee et al. [17].

3. Statistical analysis & Results

3.1 Statistical Analysis

The manuscript has developed the associations of BMI values with five explanatory variables such as age, AFC, TET, MIT and PCOS diagnosis status of the study subject. Joint generalized linear BMI values model has been obtained on the five explanatory variables such as age, AFC, TET, MIT and PCOS diagnosis status. Final BMI values model has been taken on the basis of lowest Akaike information criterion (AIC) value (within each class) that minimizes both the squared error loss and predicted additive errors [22, p. 203--204]. According to the AIC rules, JGLMs Gamma fit (AIC= 5662.777) and Log-normal fit (AIC=5663) are almost the same as the AIC difference is smaller than one. In the BMI mean model, all the included marginal and joint interaction effects are significant. Note that if any interaction effect is significant, then all its lower order interaction effects and marginal effects should be allowed in the model even if they are insignificant by marginality rule by Nelder [23]. Here all the included effects of BMI mean model are significant. Two partial marginal effects such as AGE (P=0.0839) and AFC (P=0.1170) are included in the BMI's dispersion model for better model improvement [22]. It is pointed out in Epidemiology that partial significant effects are referred as confounders that may have some influence on the risk factor or marker.

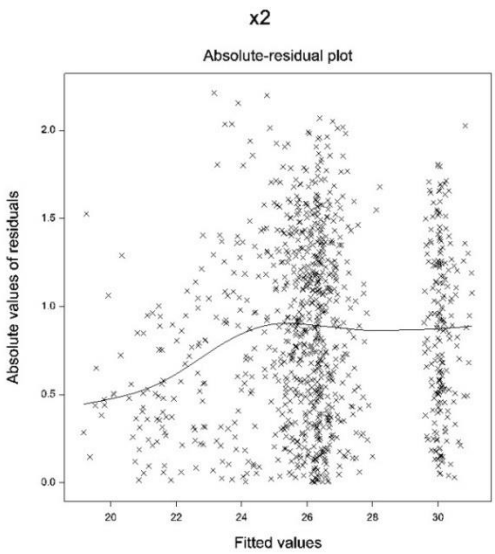

Figure 1(a)

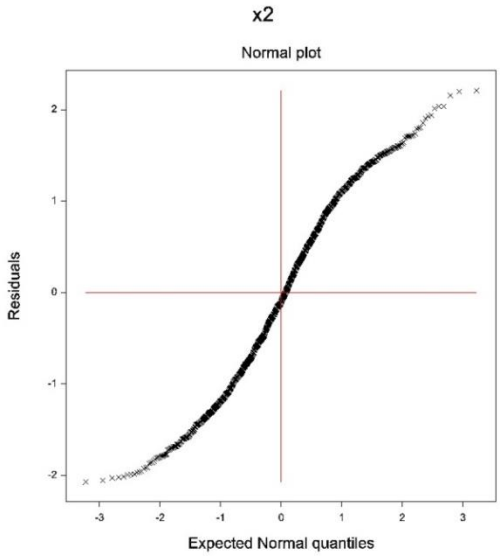

Figure 1(b)

Figure 1: For the joint Gamma fitted models of Body Mass Index (Table 1), the (a) absolute residuals plot with the fitted values, and (b) the normal probability plot for mean model

The developed BMI values Gamma fitted JGLM (Table 1) is a data extracted model that is to be tested by model checking plots. The interpretations about BMI values are taken from the data exhibited in the Gamma fitted BMI values probabilistic model (Table 1), which should be accepted based on graphical diagnostic plots in Figure 1. Figure 1(a) presents the absolute residuals plot for the Gamma fitted BMI values model (Table 1) with respect to the fitted values, which is almost flat linear, indicating that variance is constant with the running means. Figure 1(b) shows the normal probability plot for the Gamma fitted BMI values mean model (Table 1) that does not reveal any lack of fit. So, both the figures 1(a) and (1b) do not present any discrepancy in the Gamma fitted BMI values models (Table 1). The above two figures confirm that the Gamma fitted BMI values model is an approximate form of the unknown true BMI values model.

4. Results

Table 1 presents the BMI values modelling summarized results. On the basis of AIC rule, Gamma fitted (AIC= 5662.777) and Log-normal fitted (AIC=5663) JGLM show similar results for BMI values analysis. As the Gamma fitted AIC value is little lower, so the Gamma fitted outcomes are illustrated herein. These two fitted models for BMI values (Table 1) have very similar outcomes. Note that it is not always true for non-constant variance response variables. Generally, there are many discrepancies for non-constant variance response variables between the fitted Gamma and Log-normal models [24, 25].

In the current report, BMI value is considered as the response random continuous variable, and the remaining five factors such as age, AFC, TET, MIT and PCOS diagnosis status are adopted as the explanatory variables. It is developed herein that mean BMI is negatively associated with the joint interaction effect (JIE) of the subject's TET levels and MIT i.e., TET*MIT (P<0.0001), while it is positively associated with both TET (P=0.0025) and MIT (P<0.0001). Mean BMI is negatively associated with the JIE of the subject's AFC values and MIT i.e., AFC*MIT (P<0.0001), while it is positively associated with both AFC (P=0.0387) and MIT (P<0.0001). Mean BMI is negatively associated with the JIE of the subject's TET levels and AFC values i.e., TET*AFC (P=0.0253), while it is positively associated with both TET (P=0.0025) and AFC (P=0.0387). Mean BIM value is positively associated with the JIE of TET levels and the subject's PCOS diagnostic status i.e., TET*PCOS (P<0.0001), while it is positively associated with TET (P=0.0025) and negatively with PCOS (P=0.0661). Mean BIM value is positively associated with the JIE of AFC values and the subject's PCOS diagnostic status i.e., AFC*PCOS

(P<0.0001), while it is positively associated with AFC (P=0.0387) and negatively with PCOS (P=0.0661). BMI values' variance is negatively associated with age (P=0.0839), PCOS (P<0.0001) and AFC (P=0.1170).

From Table1, Gamma fitted BMI values mean ($\hat{\mu}$) model is

$\hat{\mu}$ = exp (3.1328 + 0.2671 MIT + 0.0022 TET – 0.0041 TET*MIT + 0.0048 AFC - 0.0070 AFC*MIT – 0.0969 PCOS – 0.0001 TET*AFC + 0.0032 TET*PCOS + 0.0082 AFC*PCOS),

and from Table 1, the Gamma fitted BMI values variance ($\hat{\sigma}^2$) model is

$\hat{\sigma}^2$ = exp. (- 2.986 - 0.010 AFC – 1.169 PCOS – 0.009 AGE).

From the above, BMI values mean ($\hat{\mu}$) model is explained by many marginal and interaction effects such as TET, MIT, TET*MIT, AFC, AFC*MIT, PCOS, TET*AFC, TET*PCOS, AFC*PCOS while the variance ($\hat{\sigma}^2$) model is explained by AFC, PCOS and AGE.

Model	Covariates	GAMMA FIT				LOG-NORMAL FIT			
		estimate	s.e.	t (990)	P-value	estimate	s.e.	t(990)	P-value
Mean	Constant	3.1328	0.0474	66.10	<0.0001	3.1217	0.0474	65.74	<0.0001
	MIT	0.2671	0.0462	5.78	<0.0001	0.2562	0.0463	5.534	<0.0001
	TET	0.0022	0.0007	3.03	0.0025	0.0020	0.0007	2.850	0.0045
	TET*MIT	-0.0041	0.0005	-7.59	<0.0001	-0.0039	0.0005	-7.299	<0.0001
	AFC	0.0048	0.0022	2.07	0.0387	0.0045	0.0023	1.959	0.0504
	AFC*MIT	-0.0070	0.0017	-3.98	<0.0001	-0.0067	0.0017	-3.807	0.0001
	PCOS	-0.0969	0.0527	-1.84	0.0661	-0.0699	0.0528	-1.323	0.1861
	TET*AFC	-0.0001	0.0001	-2.24	0.0253	-0.0001	0.0001	-2.049	0.0407
	TET*PCOS	0.0032	0.0005	5.57	<0.0001	0.0030	0.0005	5.291	<0.0001
	AFC*PCOS	0.0082	0.0019	4.34	<0.0001	0.0077	0.0019	4.081	<0.0001
Dispersion	Constant	-2.986	0.2013	-14.83	<0.0001	-2.979	0.2012	-14.79	<0.0001
	AFC	-0.010	0.0064	-1.56	0.1170	-0.010	0.0064	-1.530	0.1263
	PCOS	-1.169	0.1160	-10.07	<0.0001	-1.174	0.1159	-10.11	<0.0001
	AGE	-0.009	0.0052	-1.730	0.0839	-0.009	0.0052	-1.743	0.0816
	AIC	5662.777				5663			

5. Discussions

The summarized BMI values analysis results are presented in Table 1. The most appropriate BMI's mean and variance models are shown above, which are obtained from Table 1. These above BMI's mean and variance models present the associations of BMI values with the independent variables such as age, TET, MIT, PCOS diagnosis status and AFC values. The effects of BMI on PCOS women or specifically, the associations of BMI values with the factors age, TET, AFC values, MIT and PCOS diagnosis status are illustrated in the following paragraphs.

BMI's mean model (Table 1) shows that mean BMI is negatively associated with the joint interaction effect (JIE) of the subject's TET levels and MIT i.e., TET*MIT (P<0.0001), while it is positively associated with both TET (P=0.0025) and MIT (P<0.0001). This implies that BMI value increases as the joint interaction effect TET*MIT decreases. It is observed that both the marginal effects TET and MIT (0=No, 1=Yes) are positively associated with BMI values, which implies that BMI value increases for the women with MIT and higher TET levels. This is not always possible as their joint interaction effect TET*MIT is negatively associated with BMI values. So, it should not always be interpreted that women with menstrual irregularity or with higher TET levels or both may have higher BMI values. Note that if the joint interaction effect is significant, then their marginal effects are not important.

Mean BMI is negatively associated with the JIE of the subject's AFC values and MIT i.e., AFC*MIT (P<0.0001), while it is positively associated with both AFC (P=0.0387) and MIT (P<0.0001). This indicates that BMI value increases as the joint interaction effect AFC*MIT decreases. Herein both the marginal effects AFC and MIT (0=No, 1=Yes) are positively associated with BMI values, which denotes that BMI value increases for the women with MIT and higher AFC values. But it is not always possible as their joint interaction effect AFC*MIT is negatively associated with BMI values. Therefore, it should not always be concluded that women with menstrual irregularity or with higher AFC values or both may have higher BMI values.

Mean BMI is negatively associated with the JIE of the subject's TET levels and AFC values i.e., TET*AFC (P=0.0253), while it is positively associated with both TET (P=0.0025) and AFC (P=0.0387). This reveals that BMI value increases as the joint interaction effect TET*AFC decreases. Note that both the marginal effects TET and AFC are positively associated with BMI values, which denotes that BMI value increases for the women with higher TET levels, or higher AFC values, or both. But it is not always possible as their joint interaction effect TET*AFC is negatively associated with BMI values. So, it should not always be concluded that women with higher TET levels, or higher AFC values or both may have higher BMI values.

Mean BIM value is positively associated with the JIE of TET levels and the subject's PCOS diagnostic status i.e., TET*PCOS (P<0.0001), while it is positively associated with TET level (P=0.0025) and negatively with PCOS (P=0.0661) diagnosis status. This presents that BMI value increases as the joint interaction effect TET*PCOS increases. Note that one marginal effect TET level is positive and the other marginal effect PCOS (0=No, 1=Yes) diagnosis status is negatively associated with BMI values, so the joint effect TET*PCOS may not always increase. In other words, this interpretation can be restated as BMI value may be higher for the women with higher TET levels with no PCOS diagnosis status.

Mean BIM value is positively associated with the JIE of AFC values and the subject's PCOS diagnostic status i.e., AFC*PCOS (P<0.0001), while it is positively associated with AFC (P=0.0387) and negatively with PCOS (P=0.0661) diagnostic status. This shows that BMI value increases as the joint interaction effect AFC*PCOS increases. It is noted that one marginal effect AFC level is positive and the other marginal effect PCOS (0=No, 1=Yes) diagnosis status is negatively associated with BMI values, so the joint effect AFC*PCOS may not always increase. In other words, this interpretation can be restated as BMI value may be higher for the women with higher AFC values with no PCOS diagnosis status.

BMI values' variance is negatively associated with (P<0.0001). It shows that BMI values are highly scattered for the women without PCOS diagnosis status. The variance of BMI values is negatively associated with age (P=0.0839), which indicates that BMI values are highly scattered for the younger women. Also, BMI values' variance is negatively associated with AFC (P=0.1170) that shows that BMI values are highly scattered for the women with lower AFC values.

It is obtained herein that the lower joint effects TET*MIT, AFC*MIT and TET*AFC are highly risk factors for BMI values, or equivalently for PCOS women. Also, the higher joint effects TET*PCOS and AFC*PCOS are highly risk factors for BMI value. It is well-known that higher BMI value is a risk factor for PCOS women. The report has derived the associations of mean BMI value with five different joint interaction effects, and along with their marginal effects. Marginal associations of BMI values are easily understandable but the joint interaction effects are a little complex. Note that the joint interaction effects on BMI values can be located using only statistical modeling. Best of our knowledge, no earlier article identifies any joint interaction association of BMI values.

The present article has examined the associations of BMI values with the five independent variables/factors such as MIT, PCOS diagnosis status, AGE, AFC, and TET levels. There are many new concepts and problems of PCOS women that are focused in the present PCOS studies literature. The article [26] pointed out that glucose-lipid metabolism and hormonal imbalances have little results on embryo development in PCOS women. Recently an article [27] studied PCOS women that interpreted that PCOS is a complex endocrine hazard that influences 7–22% women at reproductive age groups, which can differentiate by polycystic ovarian morphology, chronic anovulation and hyper-androgenism. In addition, a recent manuscript [28] has concentrated on the reduced miR-338-3p levels that have potential predictive value in discriminating between PCOS status and normal women. Anti-Müllerian Hormone may be intertwined in regulating impaired ovarian granulosa cells improvement in PCOS rats via SMAD family member 4 (SMAD4) [29]. PCOS women with their higher BMI and hormonal factors may assail on pregnancy results such as miscarriage risk, menstrual irregularity etc. due to high androgen levels and obesity [5]. Interested researchers/practitioners may be acquainted with many new concepts of PCOS using the articles [5,10,26-30].

6. Conclusions

The present manuscript has derived the associations of BMI values with AGE, MIT, TET, PCOS diagnosis status and AFC values. The fitted BMI values model has been taken using graphical diagnostic testing plots (Figure 1), on comparison of joint Gamma and Log-normal models (Table 1), smaller standard error of the estimates and on the basis of smallest AIC rule. Both BMI value fitted Gamma and Log-normal models (Table 1) have similar interpretations based on AIC rule. All these BMI values result in Table 1 focus on the real practical situations. The derived BMI values outcomes of PCOS data set though not completely decisive but are exposing. Modern scientific research methods should have full belief on these BMI values obtained results, as the BMI values fitted models have been accepted with graphical diagnostic tools and comparing two different models. The obtained BMI value models (Table 1) are derived from the data set as mentioned in the material section. It is expected that for any PCOS data set with the same study characters, almost similar outcomes (Table 1) regarding BMI values can be obtained by any scholar that is not shown herein. The present manuscript shows many real associations of BMI values with age, AFC, MIT, TET and PCOS diagnosis status, which are not derived in the previous articles. Note that the present BMI modelling outcomes of PCOS data set are fully new in the clinical endocrine literature. It is sure that the present derived associations of BMI values with the rest five characters give fruitful information to the practitioners, researchers and common people. It concludes that women should care about their BMI values along with TET levels, antral follicle counts and menstrual irregularity.

Acknowledgement: The authors are very grateful to the principal data investigators, who provided the data freely for scientific study.

Abbreviations:

AFC: antral follicle count

BMI: body mass index

JGLMs: joint generalized linear models

JIE: joint interaction effect

MIT: menstrual irregularity

TET: Testosterone levels

PCOS: polycystic ovary syndrome

Funding: The authors declare no financial support for the research, authorship, or publication of this article.

Conflict of interest: The authors confirm that this article content has no conflict of interest.

Ethical approval: Note that the current study considers a secondary data set, which is available in the site- https://www.kaggle.com/datasets/samikshadalvi/pcos-diagnosis-dataset The ethics approval and the subject consents are not required for a secondary published data set.

Data availability statement: The data is available in the site- https://www.kaggle.com/datasets/samikshadalvi/pcos-diagnosis-dataset

Informed consent statement: Not applicable

Sample availability: The authors declare no physical samples were used in the study

References

1. Azziz, R., Carmina, E., Chen, Z., et al. Polycystic ovary syndrome. Nature Reviews Disease Primers, 2016; 2, 16057.
2. The Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group, "Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Human Reproduction, 2004; 19(1): 41–47.
3. Jabeen A, Yamini V, Rahman Amberina A, Dinesh Eshwar M, Vadakedath S, Begum GS, Kandi V. Polycystic ovarian syndrome: prevalence, predisposing factors, and awareness among adolescent and young girls of South India. Cureus. 2022, 14: e27943. 10.7759/cureus.27943
4. Neubronner SA, Indran IR, Chan YH, Pa Thu AW and Yong EL. Effect of body mass index (BMI) on phenotypic features of polycystic ovary syndrome (PCOS) in Singapore women: a prospective cross-sectional study. BMC Women's Health, 2021 21:135
5. https://doi.org/10.1186/s12905-021-01277-6
6. Wang L, Yu X, Xiong D et al. Hormonal and metabolic influences on outcomes in PCOS undergoing assisted reproduction: the role of BMI in fresh embryo transfers. BMC Pregnancy Childbirth. 2025 Mar 28;25(1):368. doi: 10.1186/s12884-025-07480-9.
7. Mohapatra I and Samantaray SR BMI and Polycystic Ovary Syndrome: Demographic Trends in Weight and Health. Cureus, 2024; 16(3): e55439. DOI 10.7759/cureus.55439.
8. Barber TM, Hanson P, Weickert MO, Franks S: Obesity and polycystic ovary syndrome: implications for pathogenesis and novel management strategies. Clin Med Insights Reprod Health. 2019, 13:10.1177/1179558119874042
9. Lim SS, Davies MJ, Norman RJ, Moran LJ: Overweight, obesity and central obesity in women with polycystic ovary

syndrome: a systematic review and meta-analysis. Hum Reprod Update. 2012, 18:618-37.10.1093/humupd/dms030

10. Zain MM, Norman RJ. Impact of obesity on female fertility and fertility treatment. Womens Health (Lond). 2008;4(2):183–94. https:// doi. org/ 10.2217/ 17455 057.4. 2. 183.
11. Pandey S, Pandey S, Maheshwari A, Bhattacharya S. The impact of female obesity on the outcome of fertility treatment. J Hum Reprod Sci.2010;3(2):62–7. https:// doi. org/ 10. 4103/ 0974-1208. 69332.
12. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod. 2016; 31:2841–55. https:// doi. org/10. 1093/ humrep/ dew218.
13. . Li Y-J, Han Y, He B. Effects of bariatric surgery on obese polycystic ovary syndrome: a systematic review and meta-analysis. Surg Obes Relat Dis. 2019;15(6):942–50. https:// doi. org/ 10. 1016/j. soard. 2019. 03. 032.
14. Lim SS, Norman RJ, Davies MJ, Moran LJ. The effect of obesity on polycystic ovary syndrome: a systematic review and meta-analysis. Obes Rev. 2013; 14:95–109. https:// doi. org/ 10. 1111/j. 1467-789X. 2012. 01053.x.
15. Indran IR, Huang Z, Khin LW, Chan JKY, Viardot-Foucault V, Yong EL. Simplified 4-item criteria for polycystic ovary syndrome: a bridge too far? Clin Endocrinol. 2018; 89:202–11. https:// doi. org/ 10. 1111/ cen. 13755.
16. Al-Muqaren HM., Mansor MB, and Ibrahim Z. Predictive modelling of polycystic ovary syndrome using machine learning. Journal of Medical Systems, 2020;.44(5): 91.
17. Nandini Modi, and Yogesh Kumar. Detection and Classification of Polycystic Ovary Syndrome Using Machine Learning-Based Approaches," 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, Gwalior, India, pp. 1-6, 2024.
18. Lee Y, Nelder JA, Pawitan Y. Generalized Linear Models with Random Effects (Unified Analysis via H–likelihood) (Second Edition). London: Chapman & Hall 2017.
19. Das RN. Robust Response Surfaces, Regression and Positive Data Analyses. London: Chapman & Hall 2014.
20. Das RN, Lee Y. Log-normal versus gamma models for analyzing data from quality-improvement experiments. Quality Engineering 2009; 21(1): 79-87.
21. Qu Y, Tan M, Rybicki L. A unified approach to estimating association measures via a joint generalized linear model for paired binary data. Communications in Statistics – Theory and Methods 2000; 29:143–156.
22. Myers, R. H., Montgomery, D. C., Vining, G. G. Generalized Linear Models with Applications in Engineering and the Sciences. New York: John Wiley & Sons, (2002).
23. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning, Springer-Verlag, 2001.
24. Nelder JA. The statistics of linear models: back to basics. Statistics and Computing 1994; 4:221–234.
25. Das RN. Discrepancy in fitting between log-normal and gamma models: An illustration. Model Assisted Statistics and Applications 2012; 7 (1) 23–32.
26. Das RN and Park JS. Discrepancy in regression estimates between Log-normal and Gamma: Some case studies. Journal of Applied Statistics 2012; 39(1): 97-111.
27. Li X., Dai J, Tang Y and Xiao T. Machine learning-based prediction models for polycystic ovary syndrome risk assessment: A review. Frontiers in Endocrinology, 2022;.13: 835477.
28. Ding K, Wang X, Liu W et al. Engineering modification of human umbilical cord mesenchymal stem cell-derived small extracellular vesicles ameliorates polycystic ovary syndrome by enhancing the ovarian environment and regulating follicular development. Stem Cell Res Ther. 2025 Sep 1;16(1):481. doi: 10.1186/s13287-025-04610-0. PMID: 40890824
29. Liang L, Lv J, Li W, Song C, Chen Y, Wei H. Reduced miR-338-3p contributes to polycystic ovarian syndrome by inhibiting proliferation and enhancing apoptosis. Hereditas. 2025 Jul 12;162(1):126. doi: 10.1186/s41065-025-00498-1. PMID: 40652291
30. Dong A, Yu X, Zhang Y, Liu L, Liu F, Song W, Zheng J. Anti-Müllerian hormone regulates ovarian granulosa cell growth in PCOS rats through SMAD4. Int J Gynaecol Obstet. 2025 Jul;170(1):209-221. doi: 10.1002/ijgo.16184. Epub 2025 Jan 24. PMID: 39865361
31. Tang T, Lai H, Huang X, Gu L, Shi H. Application of serum markers in diagnosis and staging of ovarian endometriosis. Journal of Obstetrics and Gynaecology Research. 2021; 47: 1441-1450.

DOI:10.31579/2640-1045/227