DISCRIMINANT ANALYSIS WITH LOGISTIC REGRESSION METHOD IN PREDICTION OF BABY'S WEIGHT AT BIRTH IN NIGER STATE NIGERIA

BY

GANA, Yahaya
MTech/SPS/2019/10575

DEPARTMENT OF STATISTICS
FEDERAL UNIVERSITY OF TECHNOLOGY
MINNA

SEPTEMBER, 2023

DISCRIMINANT ANALYSIS WITH LOGISTIC REGRESSION METHOD IN PREDICTION OF BABY'S WEIGHT AT BIRTH IN NIGER STATE NIGERIA

BY

GANA, Yahaya
MTech/SPS/2019/10575

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL FEDERAL UNIVERSITY OF TECHNOLOGY MINNA, NIGERIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF TECHNOLOGY IN STATISTICS

SEPTEMBER, 2023

DECLARATION

I hereby declare that this thesis titled: **"Discriminant Analysis with Logistic Regression Method in Prediction of baby's weight at birth in Niger State Nigeria"** is a collection of my original work and it has not been presented for any other qualification anywhere. Information from other sources (published or unpublished) has been duly acknowledge.

GANA, Yahaya
MTech/SPS/2019/10575
FEDERAL UNIVERSITY OF TECHNOLOGY
MINNA, NIGERIA

____
Signature/Date

CERTIFICATION

The thesis titled: **"Discriminant Analysis with Logistic Regression Method in Prediction of baby's weight at birth in Niger State Nigeria"** by GANA, Yahaya MTech/SPS/2019/10575 meets the regulations governing the award of the degree of MTech of the Federal University of Technology, Minna and it is approved for its contribution to scientific knowledge and literary presentation.

Dr A. Usman Major Supervisor	______________________ Signature & Date
Dr Y. Yakubu Co-Supervisor	______________________ Signature & Date
Dr Y. Yakubu Head of Department	______________________ Signature & Date
Prof. M. Jiya Dean, School of Physical Sciences	______________________ Signature & Date
Engr. Prof. O.K Abubakre Dean, Post Graduate School	______________________ Signature & Date

ACKNOWLEDGEMENT

All praise be to Almighty Allah the creator of universe, the lord of universe, the possession of power and authority, the fulfiller of needs, praise be to Him for giving me the opportunity, wisdom of understanding and ability to overcome this programme successfully.

First and foremost, I wish to express my profound gratitude to my Project Supervisor Dr. A. Usman for taking part of his time to go through my work and making his constructive contributions despite his tight schedule. Head of Department Dr. Y. Yakubu. My lecturers Prof. A. Isah, Dr A. Usman, Dr Y. Yisa, Dr. R. A. Adeyemi, Dr. O. M. Adetutu, Mal Abdullahi Usman, Mr Onotu Idris and Mr James Mayaki for their advice and contribution to the success of the program. I wish to express my profound gratitude to my beloved parents Mallam and Malama Muhammed Gana and my beloved wife Zainab Musa Salihu for their cares, prayers, moral support and encouragement throughout the period of my programme. To my beloved brothers and sisters Abdullahi, Suleaiman, Jibrin, Zakari, Aisha, Muhammad and Fatima. I am very grateful for their positive contributions to my education. Thank you for taking the responsibility and giving me the courage, financial support and also for taking care of me during my struggle throughout my course of study. I would like to express my appreciation to my friends Zubair Adullahi Umar, Mohammed Mahmud Sheshi and Jibrin Suleiman Danyaya, Furthermore, Finally, I wish to express my sincere gratitude to my friends for their effort and courage to see us through successfully and equally to my course mates and colleagues for their understanding throughout our period of stay in school. I wish everybody success in life and thank you all and God bless (Amen)

ABSTRACT

Low birth weight is a major public health problem in the developed and developing countries like Nigeria, contributing substantially to both infant mortality and the childhood handicap; the principal determinant of low birth weight in Nigeria is preterm birth delivery. This research work seeks to identify some factors affecting baby's weight at birth using discriminat analysis with logistic regression methods. The data were collected from medical record unit of Jummai Babangida Aliyu Maternal and Neonatal Hospital Minna Niger State. SPSS version 28.0 was used to analyze the data to see if really the factors considered affect the baby's weight at birth. The dataset meet the assumption of discriminant analysis which states that the predictors are not correlated with one another. Wilks' lambda statistics which gives 0.422 indicates greater discriminatory ability of the function. From the selected cases 168 of 206 i.e (80.7%) of low birth weight were correctly classified and 40 (19.2%) were misclassified. 337 of 396 i.e (85.1%) of normal birth weight were correctly classified while 59 (14.9%) were misclassified. Among the seven maternal characteristics examined, the parameter estimate of the model using Wald's statistic, two variables maternal height and gestational age are statistically significant while the remaining are insignificant at $\alpha = 0.05$. It has been shown that increase in mothers weight by 1kg brings about some increase in baby's weight. It is recommended for further research to see if there are other maternal characteristics that brings about effect on baby's weight at birth. The model build by this study should be adopted to discover the prevalence of low birth weight among infants so that adequate measures for prevention and control of birth weight can be taken early enough.

TABLE OF CONTENTS

Contents **Page**

Title page		i
Declaration		ii
Certification		iii
Acknowledgement		iv
Abstract		v
Table of content		vi
List of Tables		viii
CHAPTER ONE		
1.0	**INTRODUCTION**	1
1.1	Background to the study	1
1.2	Statement of the Research Problem	15
1.3	Aim and objectives of the study	15
1.4	Justification for the study	15
1.5	Scope and Limitation of the study	16
1.6	Definition of terms	16
CHAPTER TWO		
2.0	**LITERATURE REVIEW**	18
2,1	Empirical framework	18
2.1.1	Conceptual framework	20
2.1.2	Theoretical framework	26
2.2	Gap identified in the Literature	32
CHAPTER THREE		
3.0	**MATERIALS AND METHODS**	33
3.1	Study Design	34
3.2	Population of the Study	34
3.3	Data collection	34
3.4	Discriminant Analysis Procedure	34
3.4.1	Method used in the study	36

3.4.2 Test for significance of canonical correlation (Wilk's lambda) 40
3.4.3 Omnibus chi-square test 42
3.4.4 Box's M-Test for the equality of covariance matrices 43
3.4.5 Logistic regression model 44
3.5 Wald's Test 47
3.6 Hosmer Lemeshow's Goodness of Fit 48
3.7 Confusion Matrix 49
3.8 Multicollinearity 50
3.8.1 Effects/Consequences of multicollinearity 50
CHAPTER FOUR
4.0 **RESULTS AND DISCUSSIONS** ` 51
4.1 Preamble 51
4.2 Presentation and Discussion of Results 51
4.3 Linear Discriminant Function 54
4.4 The Constructed Logistic Regression Model 59
4.5 Hosmer and Lemeshow test 62
4.6 Checking for Multicollinearity 62
4.7 Major Findings 63
CHAPTER FIVE
5.0 CONCLUSION AND RECOMMENDATIONS 64
5.1 Conclusion 64
5.2 Recommendations 65
5.3 Contribution to knowledge 65
REFERENCES 66
APPENDIX 70

LIST OF TABLES

Table		page
4.1	Descriptive Statistics	51
4.2	Correlational Matrices of the indicator variables (Pooled within-Groups)	52
4.3	Pooled covariance Matrices	52
4.4	Test of Equality of Group Means	53
4.5	Test Results of Box's M	53
4.6	Wilk's Lambda Test	54
4.7	Fisher's Classification Function Coefficient	55
4.8	Standardized Canonical Discriminant Function Coefficients	57
4.9	Functions at Group Centroids	58
4.10	Prior Probability for Groups	58
4.11	Classification Results	59
4.12	Variables in the Equation for the sample Data	60
4.13	Testing for Multicollinearity	62

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to the Study

Birth weight is an important indicator for assessing future growth patterns of children and investigating direct health risks and in adulthood. It is a key variable in any longitudinal study of child health.

It is the weight of the child at birth and be defined as the baby's first weight gain in the first 60 minutes after birth. High birth weight refers to babies weighing more than 4kg, full-size or normal babies weighing between 2.5kg and 4kg, low birth weight (LBW) means babies weighing less than 2.5kg, and low weight at birth (i.e less than 2.5kg) increases prenatal death rates, causes range from premature birth to placental insufficiency. During pregnancy, the baby's birth weight can be estimated in different ways and the height of the fundus (the top of the mother's uterus) can be measured from the public bone. This measurement in centimeters generally corresponds to the number of weeks pregnant after week 20. The clinical manifestations of babies at birth vary because many factors affect the size of the baby at birth. A little bit of this not a problem, and a lower weight does not always indicate stunted growth. A woman's first child is usually lighter than her later children, and mothers who are very old at birth often give birth to older children. The size of the baby's father also affects his birth weight.

Birth weight is the first weight of the fetus or newborn obtained after birth, preferably measured within the first hour of life before significant postnatal weight loss has occurred. Low birth weight (LBW) by international agreement has been defined as a birth weight of less than 2.5kg. World Health Organization (WHO 2020). According to the WHO's estimate, the global rate of LBW in 2018 was 15.5 percent and the rate in developing countries (16.5 percent) was more than double that of developed countries (7.0 percent).

Low birth weight babies are the result of premature delivery (before 37 weeks of gestation) or fetal growth restriction (in uterus) within one week of pregnancy. Low birth weight is closely related to the mortality and morbidity of fetuses and newborns. It inhibits growth and cognitive development, as well as chronic diseases of the elderly. There are many factors that affect the length of pregnancy and fetal development, including birth weight. They are related to the baby, the mother, or the physical environment. They play an important role in determining weight at birth and beyond. In facts some babies born short are prone to certain health problems. Some of you may be sick or infected from the first day of life. Others may have long-term problems, such as movement and social delays or learning disabilities. Low birth weight babies are usually tall because their father is tall or their mother has diabetes during pregnancy. The risk of birth of these babies may be higher. Trauma and blood sugar problems. Also according to WHO and United Nation Children Fund (UNICEF) (2004). Birth weight depends largely on the mother herself. The development of the fetus and its upbringing from birth to pregnancy, and its body. The component to be designed. Mothers with difficult socioeconomic conditions often give birth to low weight babies. It was also found that wealthy women had better outcomes in childbirth. Therefore, the mother's economic level is the most important factor that determines the child's living standards. Height at birth. This is because prenatal care is better than that of the wealthiest women. Premature babies are at increased risk of the following diseases. Developmental disabilities are more important than full-term babies. The first difficulty is that the quality of interaction between parents and children is very poor.

Prenatal and perinatal complications in preterm infants (such as neonatal cerebral hemorrhage) may pose potential risks to brain development. Neuro developmental skills later in life.

Low birth weight is a major public health problem in developing countries including Nigeria. Epidemiological observations have shown that infants weighing less than 2.5kg are approximately 20 times more likely to die than infants heavier, closely related to fetal morbidity and mortality and Infant. In India, 3.035 percent of babies are born with low birth weight and more than half of them are full-term. Low weight loss is an important contributor to the Millennium Development Goals (MDGs) to reduce child mortality. Actions to achieve the MDGs will need to ensure children have a good start in life by ensuring that women start their pregnancies healthy, are well nourished, and get through pregnancy and childbirth safely. Understanding the prevalence and factors contributing to and perpetuating this problem will help address this important cause of infant mortality in order to reduce and achieve the Millennium Development Goals. Identifying the factors that cause the still high rate of low birth weight and putting in place remedial measures to combat this problem should be seen as a major public health challenge. In this context, the study was carried out to find out the prevalence of low birth weight in term infants and to discover maternal risk factors associated with low birth weight infants.

Birth weight or height is an important indicator of a child's vulnerability to childhood disease risks and can predict a child's health, development, and chance of survival. Children in the future. The WHO (2004) defines low birth weight (LBW) as an infant weighing less than 2.5kg. This is based on epidemiological observations that infants weighing less than 2.5kg have a higher risk of neonatal mortality compared with heavier infants. Low birth weight is considered the most important predictor of infant mortality, especially in the first months of life. Globally, between 60% and 80% of infant deaths occur in low birth weight infants. In developing countries, a birth weight of less than 2.5kg is a major cause of infant mortality and causes many health problems. It is associated with cognitive and neuro developmental dysfunction, childhood morbidity, growth retardation, various adverse health effects, and chronic illnesses later in life. It is responsible for the short and long term consequences of negative economic and social impacts.

Globally, more than 20 million babies are born with a low birth weight, accounting for 15.5 percent of all births; 95.6 percent of them live in developing countries, accounting for 17 percent of all births in developing countries. According to a study of childbirth in a hospital in Iran, the rate of children with low birth weight is 40 percent, gestational age less than 37 weeks, maternal age less than 20 years, uneven prenatal check-up, a maternal size. less than 150 cm and the mother's weight less than 50 kg, hemoglobin less than 10 g/dl, heavy physical labor and chewing tobacco are important factors determining low birth weight. Epidemiological investigations in China revealed low birth weight and maternal age under 20 years, low maternal education, poor pregnancy history, and pregnancy complications and complications (such as 'increased') blood pressure during pregnancy, anemia, premature rupture of membranes and gestational diabetes. In a study in Indonesia, determinants of low birth weight included baby's sex, woman's education level, season of birth, mother's place of residence, family assets, mother's height, birth order, pregnancy interval. According to the WHO, factors leading to low birth weight in developing countries include inadequate weight gain during pregnancy, low birth weight before pregnancy, short stature, infectious diseases such as malaria, heavy manual labor during pregnancy, and social factors such as the low status of women. Malnutrition and lack of prenatal care. In addition to regional variation, low birth weight is more common among young mothers (under 20 years of age) and children of older mothers. Furthermore, single child births, children of uneducated mothers, and mothers in the richest quintile are often reported to be very young. On the other hand, various studies report that factors such as lack of prenatal care, premature births, chronic diseases, lack of formal education and young mothers are associated with low birth weight infants. It is important to understand the risk factors of low birth weight infants in order to properly identify and manage mothers at risk. Although there are few national studies that indicate factors associated with low birth weight, they were not studied in the study area. Therefore, this study examined factors associated with low birth weight, particularly those associated with childbirth at the Debreberhan Referral Hospital. The results of this study enhance our current understanding of the risk factors for low birth weight, particularly in terms of method of delivery, chronic maternal diabetes, pregnancy complications, and trauma. Physical during pregnancy without our knowledge. This also explains further research. This has important implications in identifying mothers and children at risk, designing appropriate measures and prompting interventions.

It was also understood that high birth weight may be regarded as a predictor for dental caries, and especially, birth weight ≥ 4.5kg is a risk factor for caries increment during adolescence another result indicated positive and significant relationship between high birth weight and bone tumour risk. Further individual with high birth weight were found to be more likely to develop, while European population with high birth weight exhibited a greater risk for bone tumor. The interest of the study lies on the immediate factors that affect the babies weight, the weight expressed in kilograms as its units is obviously dependent on many factors but for this piece of work our main objective is to check if mothers weight, gestation period, age, parity and sex of a baby is in any way related to the weight of baby while giving birth to the baby.

Thanks to planned efforts within the framework of the Millennium Development Goals (MDG's), the burden of infant mortality has been significantly reduced in the past few decades. From 1990 to 2013, the global mortality rate of children under 5 years of age was approximately halved (90 percent in 1990 and 46 percent per 1,000 live births in 2013). Despite progress in total infant mortality, the neonatal mortality rate continues to rise (38% in 2000 and 45% in 2015), posing a major obstacle to achieving the Millennium Development Goals. Globally, premature birth (28 percent), severe infection (26 percent) and asphyxia (23percent) are the most important causes of neonatal death. However, low birth weight (LBW) (<2.5kg at birth) is also considered an important underlying determinant and a contributing factor in neonatal and infant mortality. Low birth weight accounts for almost half of all perinatal deaths and one third of all infant deaths. Compared to normal birth weight (NBW) babies, babies with LBW are 40 times more likely to die within the first 30 days of life. In African countries, low birth weight is considered the strongest predictor of infant morbidity and mortality. In view of the importance of LBW for child survival, the 34th General Assembly of the WHO (2020) adopted LBW as one of many health indicators as part of the global health strategy. Regional statistics show that the global burden of neonatal mortality is heavily skewed towards low and middle-income countries which account for almost all LBW cases. According to WHO estimates, of the more than 200 million low-birth-weight babies (15.5 percent of all live births), almost 95.6 percent are in low and middle-income countries.

Linear discriminant analysis and logistic regression method are multivariate statistical methods, and they are the two most commonly used methods. A popular method for solving classification problems involving binary categorical variables Yarnold *et al*., (1994).

Logistic regression predicts the probability of belonging to a group related to multiple variables it has nothing to do with its distribution. Logistic regression is based on calculating the probability of obtaining a result divide by the probability of not having it. Logistic regression is not parameterized, assuming free distribution Performance. On the other hand, discriminant analysis is used to determine which set of variables can be distinguished between two or more natural groups, and classify observations into these known groups. It is the parametric method assumes that the sample comes from a normally distributed population, and the covariance matrix of the independent variables is the same for all groups. Several authors formally compared the two technologies..
Dattalo (1995) found that the two methods work well as a classification technique, but the conclusion is that logistic it is more concise. Discriminant analysis helps to classify observations from one of the two groups, while Logistic regression helps to associate a qualitative (binary) dependent variable with one or more independent variables.

Montgomery *et al*., (1987) cited Kleinbaum *et al*., (1982) compared the classification ability of the two methods using the data set that satisfies the hypothesis of discriminant analysis and indicates this logistic regression model is slightly better.

Edokpayi *et al., (*2013) compared the set two methods in Classify and evaluate the relative importance of fruit shape features, but the conclusion is that both the value of these methods is almost the same, but as long as normality, logistic regression will be desirable Assumption is violated. Balogun *et al., (*2014) compared two methods for classifying and evaluating the characteristics of drug offenders, but The bottom line is that these two methods provide a close value, but logistic regression will be desirable as long as Violation of the assumption of normality. Based on the previous arguments, the purpose of this study is to compare these two analysis methods using data sets. Payment method, Like any other model building technique, the goal of the logistic regression analysis is to find the best fitting and most parsimonious, describe the relationship between an outcome (dependent or response variable) and a set of independent. Hosmer and Lemeshow (1989). This statement motivates the purpose of this study: to identify risk factors for low birth weight (LBW) in newborn infants using the statistical tool of logistic regression analysis. For example to determine the risk factors for low birth weight, data could be collected on several variables, such as weight of child at birth, gestational age, child sex, etc. The response variable here, is dichotomous that either the weight of child at birth is low (Y = 1), or is not low (i.e., normal) (Y = 0). Thus Logistic Regression is a mathematical modeling approach that can be used to describe the relationship of several independent variables to (say) a binary (dichotomous) dependent variable.

In clinical situations, the status of a patient is assessed by the presence or absence of a disease. There are many factors to consider which may or may not correlate with the incidence of the disease. There has been numerous retrospective medical research studies published each year that review past medical records and charts of former patients to help determine some of the risk factors (or causing agents) of diseases that are of interest. Finding the risk factors and the potential risk factors can help prevent the development of the disease. All of the diseases and nearly all of the risk factors considered are categorical variables (variables taking on two or more possible values). Hosmer and Lemeshow (1989), two prominent statisticians, state that “the logistic regression model has become the standard method of analysis in this situation.”

In modern days, statistics has played a significant role in biological, pharmaceutical and medical sciences. Cornfield (2010). The application of multivariate statistical techniques to biological and medical data has dominated the areas of evidence-based medicine. Multivariate methods are relevant in virtually every branch of applied medicine, pharmacy and public health. They come into play either when we have a medical theory to test or when we have a relationship in mind that has some importance for medical decision or policy analysis in public health.

Interest in human development before birth is widely spread because of the interest in knowing more about our beginning and the desire to improve the quality of life. The intricate process by which a baby develops from a single cell is miraculous and few events are more exciting than a mother's viewing of her embryo during an Ultrasound examination. Human development is a continuous process that begins when an ocyt (ovum) from a female is fertilized by a sperm (spermatozoa) from the male. By accepting the shelter of uterus, the foetus also takes the risk of material disease or malnutrition and of biochemical immunological and hormonal adjustment.

Until the beginning of the nineteenth century, far more attention was paid to the collection and presentation of data than to their interpretation. Large volume of data were usually collected and frequently misinterpreted if indeed interpretation was attempted. However, since that time, the importance of a scientific approach in the interpretation of data has been realized and great steps have been achieved in the development of appropriate methods.

Multivariate methods are prominently used on data to test a theory or to estimate a relationship in medicine, pharmacy and public health. In some cases, especially those that involve the testing of medical theories, a formal multivariate model is constructed. The model consists of multivariate technique that describes various relationships. In most cases, the model is used to make predictions in either the testing of a medical theory or the study of a policy impact in pharmacy and public health.

Kirkwood and Stern (2008) defined Discriminant analysis and classification as the multivariate techniques concerned with separating sets of objects or observations and with allocating new objects or observations to previously defined groups. As a separation procedure, it is often employed on a one-time basis in order to investigate observed differences when causal relationships are not well understood. The immediate goals of Discriminant analysis and classification are to describe the differential features of objects so as to find discriminant whose numerical values are such that the collections are separated as much as possible and to sort new objects or observations into two or more classes or groups.

The purpose of discriminant analysis is to correctly classify observations or people into homogeneous groups. The independent variables must be metric and must have a high degree of normality. Discriminant analysis builds a linear discriminant function, which can then be used to classify the observations. The overall fit is assessed by looking at the degree to which the group means differ (Wilk's Lambda) and how well the model classifies. To determine which variables have the most impact on the discriminant function, it is possible to look at partial F values. The higher the partial F, the more impact that variable has on the discriminant function. This tool helps categorize people, like buyers and non-buyers.

Discriminant analysis is used to build Discriminant functions which are linear functions of variables that can be used to describe or elucidate the differences among $p \geq 2$ groups. The goals of discriminant analysis include identifying the relative contribution of the p variables to separation of the groups and finding the optimal plane on which the points can be projected to best illustrate the configuration of the groups. Another goal of discriminant analysis is the prediction or allocation of observations to groups, in which linear functions of the variables are employed to assign an individual sampling unit to one of the groups. The measured values in the observation vector for an individual or object are evaluated by the classification function to find the particular group, to which the individual most likely belongs.

In clinical situations, the status of a patient is assessed by the presence or absence of a disease. There are many factors to consider which may or may not correlate with the incidence of the disease. There has been numerous retrospective medical research studies published each year that review past medical records and charts of former patients to help determine some of the risk factors (or causing agents) of diseases that are of interest. Finding the risk factors and the potential risk factors can help to prevent the development of the disease. All of the diseases nearly all of the risk factors considered are categorical variables (variables taking on two or more possible values). Hosmer and lemeshow (1989), two prominent statisticians, state that ‘the logistic regression model has become the standard method of analysis in this situation.

Like any other model building technique, the goal of the logistic regression analysis is “to find the best fitting and most parsimonious, yet biologically reasonable model to describe the relationship between an outcome (dependent or response variable) and a set of independent (predictor or explanatory) variables”, Hosmer and Lemeshow (1989). This statement motivates the purpose of this study: to identify risk factors for low birth weight (LBW) in newborn infants using the statistical tool of logistic regression analysis. For example to determine the risk factors for low birth weight, data could be collected on several variables, such as weight of child at birth, gestational age, child sex, etc. The response variable here, is dichotomous that either the weight of child at birth is low (Y=1), or is not low (i.e., normal) (Y=0). In such cases, the usual MLR theory is not appropriate. Rather, the statistical model preferred for the analysis of such binary (dichotomous) responses is the binary logistic regression model. Thus Logistic Regression is a mathematical modeling approach that can be used to describe the relationship of several independent variables to (say) a binary (dichotomous) dependent variable.

The use of logistic regression first appeared during the mathematical studies for the population growth at that time. The term logistic regression analysis comes from logit transformation, which is applied to the dependent variable. This case, at the same time, causes certain differences both in estimation and interpretation

Logistic regression analysis is also called "Binary Logistic Regression Analysis", "Multinomial Logistic Regression Analysis" and "Ordinal Logistic Regression Analysis" depending on the scale type where the dependent variable is measured and the number of categories of the dependent variable. Logistic regression is divided in to two; Univariate Logistic Regression and Multivariate Logistic Regression.

Logistic regression sometimes referred to as "choice models," this technique is a variation of multiple regressions that allows for the prediction of an event. It is allowable to utilize non metric (typically binary) dependent variables, as the objective is to arrive at a probabilistic assessment of a binary choice. The independent variables can be either discrete or continuous. A contingency table is produced, which shows the classification of observations as to whether the observed and predicted events match. The sum of events that were predicted to occur which actually did occur and the events that were predicted not to occur which actually did not occur, divided by the total number of events, is a measure of the effectiveness of the model. This tool helps predict the choices consumers might make when presented with alternatives.

Logistic regression analysis (LRA) extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical. In practice, situation involving categorical outcomes are quite common. In the setting of evaluating an educational program, for example, predictions may be made for the dichotomous of success/failure or improved/not-improved. Similarly, in a medical setting, an outcome might be presence/absence of disease. The focus of this record is on situations in which the outcome variable is dichotomous, although extension of the techniques of LRA to outcomes with three or more categories (e.g improved, same, or worse) is possible.

In many application areas, such as epidemiologic and biomedical studies, where outcomes may be occurrence or nonoccurrence, mortality (dead or alive), and so forth, logistic regression is the standard approach for the analysis of binary and categorical outcome data.

In this era of information and communication technology, as well as the era of evidence-based medicine, statistical modeling has become as necessary the medical practitioners who are interested in lasting solution to diagnosed problems. In this twenty first century, statistics play an important role in many simulations, modeling and decision-making processes in the medical sciences. This implies the need for statistical research in every facet of medicine; especially the evidence-based medicine. It was mentioned that the critical factor that separates statistical research from other ways of knowing the medical world is that statistical research is purely scientific in nature. In this sense, Science refers to both a system for producing medical knowledge and the knowledge produced. Also Science is a combination of an orientation towards a set of procedures, techniques, knowledge and instruments for gaining knowledge.

This research work is to determine if there is convergence between the two analytical methods Classification of subjects (birth weight babies) into one of two groups (Low birth weight) and (Normal birth weight), and confirmed the validity of the assumptions on which the two methods are based. In choosing between these two methods, the study shall apply the following criteria, namely the group prediction Membership and evaluation of its success, that is, determining which of the two methods provides a greater. The accuracy of the classification of birth weight for pregnant women. Determine what variables are present It is important to classify the dependent variables by checking coefficients and testing hypotheses. Normality and equal covariance required for the validity of the discriminant analysis.

The results of the outcome of the study shall not only complement the current practices but will also assist the research scientists to make appropriate choice in their application of these two techniques.

1.2 Statement of the Research Problem

Low birth weight is a major public health problem in the developed and developing countries like Nigeria, contributing substantially to both infant mortality and the childhood handicap; the principal determinant of low birth weight in Nigeria is preterm birth delivery, a phenomenon of largely unknown etiology. Preterm delivery is more common in Nigeria than in any other industrialized nations.

1.3 Aim and objectives of the Study

The aim of the study is to set up discriminant function with logistic regression method in prediction of baby's weight at birth

Specific objectives are to:

i. determine if there is convergence between the two methods of analysis classifying the birth weight into one of the two populations (Low birth weight and Normal birth weight).

ii. determine the tenability of the assumption underlying the two methods.

iii. estimate the probability of correct classifications and misclassifications respectively.

iv. determine which variables is significant in classifying the dependent variable by inspection of the coefficient.

1.4 Justification of the study

In Nigeria there is no recognized scientific method of discriminating and classifying babies statistically into groups of study. This has prompted the researcher to set up the scientific method to model for classification. The study is necessary because it will assist a medical researcher to ascertain the prevalence of these factors in a given population.

1.5 Scope and limitation of the study

The study is limited to the data collected and will only focus on the sample taken from Jummai Babangida Aliyu Maternal and Neonatal Hospital Minna Niger State. This would be used for the analysis among women.

The study build the models on seven predictor variables; i.e maternal height, maternal age, baby's weight, baby's weight, baby's sex, gestational age and Parity. Only seven predictor variables incorporated in both the linear discriminant and logistic regression models as the most relevant factors considered and captured by the study.

1.6 Definition of terms

i. **Low Birth Weight (LBW)** is described as a birth weight of a live born infant of less than 2.5kg regardless of gestational age.

ii. **Gestational age**: the weeks or months of pregnancy. A normal pregnancy can range from 38 to 42 weeks.

iii. **Parity:** the number of pregnancies reaching viable gestational age including live births and still births.

iv. **Weight**: this is how heavy somebody or something is which can be measured in kilograms.

v. **Height:** this is the measurement of how tall a person or thing is and it is measured in metres

vi. **Discriminant analysis (DA)**: are multivariate techniques concerned with separating distinct sets of objects (or observations) and with allocating new objects (or observations) to previously defined groups.

vii. **Logistic regression:** deals with the binary cases, where the response variable consists of just two categorical values; it is mainly used to identify the relationship between two or more explanatory variables X_i and the dependent variable Y.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Empirical Review

Pathak *et al.* (2004) in a univariate and multivariate analysis study revealed that no variable was significantly associated with the copper deficiency and birth weight.

However, a significant increase in the serum copper levels was found with the increase in pregnancy duration. In another observational study in Poland,

Higgins *et al.* (1989) reported an analysis of "Higgins Nutritional Intervention program" participants from the "Montreal Diet Dispensary" program, which were high risk mothers from a nutritional standpoint and managed by a specific nutritional rehabilitation program depending upon need. An analysis of 525 mothers who participated in their second pregnancy but not their first pregnancy was reported. The comparison was made with birth weight of the first and the second child and it was observed that the rate of LBW births was lower (4.9 percent vs. 8.9 percent), the mean birth weight was 107g higher in the intervention group ($p < 0.01$) and the rate of IUGR births was lower (1.4 percent vs. 2.4 percent) among participants. The authors concluded that there was a benefit of the intervention among low-income high-risk women. This effect may represent a natural phenomenon, as second born infants are usually heavier.

Wasowicz *et al.* (1993) found that both zinc and copper concentrations in plasma of preterm infants were significantly higher than in full term infants. Mothers of preterm infants did not differ in plasma zinc and selenium levels but copper concentrations were significantly higher as compared to mothers of full term neonates. Mothers' giving birth to low birth weight babies has significantly higher copper levels as compared to those giving birth to heaviest babies. Similarly, serum copper measured at delivery was associated negatively with birth weight.

Herrera *et al,* (2020) developed an antenatal bio psychosocial risk assessment tool which was found to be very effective at prediction of low birth weight in a double-blind prospective study in USA. This study included 1,076 mother-infant pairs and assessed bio psychosocial risk as perceived by the mother on factors including reproductive history, medical history, anxiety and concerns about financial situation. The authors concluded that the tool should be further trailed among larger sample groups. However, this study supported others in the case for mandatory conclusion of social factors when considering issues of low birth weight babies.

Kramer (2022) reviewed the impact of supplementation of a balanced protein/energy diet (where the protein content of diet was < 25 percent of the total energy content) on gestational weight gain and pregnancy outcomes from 13 studies for the Cochrane Collaboration. The quality of the trials varied and often the method of randomization were not stated. It was found that there was an increase in maternal weight gain (17g/week) and a reduction in the risk for of SGA births. There was no difference in the stratified analysis of undernourished (determined based on pregnancy weight) and adequately nourished women in terms of difference in birth weight with supplementation of adequate nutrition (24g vs. 25g). No difference was found in the risk of preterm births.

Rees *et al,* (2005) found lifestyle behaviors such as inadequate nutrition, smoking mothers themselves who were low birth weight, low pregnancy weight gain, increasing maternal stress and / or depression, domestic violence and maternal regret and/ or rejection of pregnancy to be significant factors. Other socio-demographic factors were low maternal age (under 18), high maternal age (over 35), low educational level, poverty, ethnicity and late or no antenatal care.

Adler and Donlon (2020) in the study concluded that morphometric crown traits in the deciduous dentition can be used to classify sex of juvenile skeletons (11 months to 12 years) of European descent from linear discriminant functional analysis with accuracy between 70.2 percent and 74.8 percent.

2.1.1 Conceptual framework

Fernandez *et al.,* (2021) used a discriminant analysis to investigate whether FT – Raman spectroscopy as spectroscopic fingerprint techniques combined with some chemometric tools can be used as a rapid and reliable method for the discrimination of honey according to their sources. In their study, they used developed models exploratory techniques as the fishers criterion, supervised methods as partial least squares -discount analysis (PLS-DA) or support vector machine (SVM) which all show a correct classification ratio between 85 percent and 90 percent of average showing Raman spectroscopy combined to chemometric treatment is a promising way for rapid and non-expensive discrimination of honey according to their regions.

Ngwu (2009) in the study of birth weight of babies in relation to their nutrition knowledge and place of residence found (12 percent) prevalence of LBW similar to national average of 12 percent and to 12.64 percent, 11.4 percent and 12.6 percent reported elsewhere. This is much higher than 8.2 percent reported by reflecting the worsening economic situation in the present day Nigeria and a frightening future trend in the face of unabated current global food crises and economic meltdown.

Onyiriuka (2018) discovered that maternal parity have a significant influence on the incidence of delivery of LBW infants in twins gestations. As in previous studies, the incidence was higher in primiparous compared with multiparous counterparts, suggesting that the uteri of multiparous women are more efficient in nurturing and promoting intrauterine growth of twins; accounting for the relatively lower incidence of LBW twin infants among them.

Eneh and Ugwu (2021) in a similar study on proportion of low birth weight babies due to small for gestational age revealed that the incidence of LBW in south – south Nigeria was 10.1 percent similar to the incidence of 10.31 percent in Enugu, but lower than the

19.8 percent reported in Kano city (North West Nigeria). In south West Nigeria, the incidences ranged from 8.2 percent to 16.8 percent while in Plateau (North Central Nigeria), the incidence was 12.2 percent. These values are in conformity with the WHO estimates that the low birth weight levels in majority of sub-Saharan Africa fall between 10% and 20 percent.

In twin gestation, prematurity is a more important contributor to delivery of LBW infant than term SGA. The finding strongly challenges the report of some studies which stated that majority of LBW infants in developing countries were due to term SGA rather than preterm delivery. The risk of prenatal death was higher in very preterm infants, suggesting that there is a need to make effort to prolong gestation beyond 32 weeks. One way of achieving this, is by instituting a prophylactic hospitalization policy (bed rest) for all women with twin pregnancy between 28 and 32 weeks gestation.

Beki (2012) used multivariate discriminant analysis and binary logistic regression for tracking the incidence of Broncho-Pulmonary Dysplasia among infants. Hence, the researcher used three possible predictor variables i.e. weight at birth, weight four weeks later and gender and built a discriminant model that is capable of tracking Broncho-pulmonary dysplasia (BPD) infants

The study predicted the BPD status of five new infants using the discriminant model in which all the five new cases were correctly predicted. Hence, the discriminant model built has a perfect classification of five new cases in Kaduna while it has misclassification of one of five new cases in Sokoto. Conversely; the study has predicted the BPD status of five new infants using logistic model in which all the five new cases were correctly predicted or classified. Hence, the logistic model built has a perfect classification of five new cases in Sokoto while it has misclassification two of five new cases in Kaduna. The probability of the classification in the study indicates that, the prior probability of misclassifying health infants to BPD is 0.05 and prior probability of misclassifying BPD infants to health is also 0.05.

Danbaba *et al.* (2013) in the study carried out work on Low birth weight using multivariate logistics regression analysis to determine the prevalence of low birth weight (LBW) and some of its risk factors in maternity hospitals in wushishi local government in Niger State. A sample of 200 live births was collected in the hospital from June-September 2011. The data were collected by measuring the mother's age at birth, mother's weight at birth, mother's education level, mother's occupation, gestational age, birth interval, twin or singleton birth and parity. The study fit the logistic regression model to the data. The analysis of variance and chi-square tests were used to know the variables of factors that have statistical significance effect on birth weight at 95% confidence level. The odds ratio (OR) of the risk factors of LBW was found using a multivariate logistic regression. They established the fundamental model for multiple regression analysis, with the assumption that the outcome variable was a linear combination of a set of predictors. For outcome variable Y, and a set of n- predictor variables $X_{1,}$,$X_{2,}$…$X_{n,}$ used the model as; Where is the expected value of y when x's are set to 0, is the regression coefficient for each corresponding predictor variables, and e is the error of the predictor.

The analysis shows that there is no significance in prevalence between boys and girls (14.9 percent versus 13.9 percent) i.e p = 0.578

Vishwa *et al.* (2015) stated that discriminant analysis and classification are multivariate techniques concern with separating district sets of objects (or observations) and with allocating new objects (or observations) to previously defined groups. Discriminant analysis is rather exploratory in nature. As a classificatory procedure, it is often employed on a one-time basis in order to investigate observed differences when causal relationships are not well understood. Classification procedures are less exploratory in the sense that they lead to well-defined rules, which can be used for assigning new objects. It is possible to have classifications into two or more multivariate normal populations, but the study shall be restricted to classifications into two normal populations denoted by π_1 and π_2.

The majority of LBW in developing countries were initially said to be due to intrauterine growth retardation. This was collaborated by earlier studies. The proportion of preterm among the low birth weight population was 75.4 percent. Other studies have also noted that the majority of the LBW were mainly contributed by prematurity. In Enugu, 69.05 percent and in Plateau 61 percent of the LBW were preterm. This denotes a changing trend from initial reports that suggest that LBW in developing countries are due predominantly to SGA.

Birth weight is an important predictor of infant survival. Infant born with a low birth weight tend to have extremely high rate of morbidity and mortality. This was clearly demonstrated in the study were the mortality among the ELBW population was above 90 percent where as it was 18% for those infants weighing between 2500g and 3999g and 7.1 5 for those weighing greater than 4000g. LBW is associated with impaired immune function which increases mortality from infectious diseases.

Uthman (2018) effect of birth weight on infant mortality found that children born with low birth weight are more likely to die during the first year of life compared to children born with normal weight, independent of child's sex, birth order, pregnancy care and delivery care, maternal education and nutritional status, household access to clean water and sanitation, and other factors.

Erimafa *et al*. (2009) used discriminant analysis to predict the class of degree obtainable in a University system. In this study, it was clearly stated that, the conditions for predictive discriminant analysis were obtained, and the analysis yielded a linear discriminant function which successfully classified or predicted 87.5 percent of the graduating students' class of degrees. The function had a hit ratio of 88.2 percent when generalized.

Mobil *et al.* (2020) used both principal component Analysis (PCA) and partial least square– discriminant analysis (PLS-DA) to analysis and interpretation of the Raman spectra collected from microorganism of different species recorded in the spectral range of 2000 to 200 cm^{-1}. To develop a classification rule, the researcher used PLS-DA in a LOOCV method for the calibration and validation of a classification model. It was asserted that, results obtained showed an acceptable classification among the strains under study; thereby, suggested it to be useful tools for the classification and discrimination of similar samples.

Rees *et al.* (2005) found lifestyle behaviors such as inadequate nutrition, smoking, mothers themselves who were low birth weight, low pregnancy weight gain, increasing maternal stress and/or depression, domestic violence and maternal regret and/or rejection of pregnancy to be significant factors.

Similarly, in agreement with other studies, Abalkhail and Khalid (1995) confirmed that fetal weight at birth influenced by, besides the mother's health status, a variety of biological, social and even geographical factors. Most of these factors are known to have variable prevalence in different regions even in the same country. An example of this is the study in Taif city, which differs from other areas in Saudi Arabia in being at high altitude. The incidence of LBW was almost double that reported from other areas of Saudi Arabia. Therefore, it may, neither clinically nor epidemiologically, be appropriate to apply the WHO cut-off level of 2.5kg for identifying LBW infants in the local population. Indeed, it has been suggested that for ethnically homogenous populations there are fundamentally normal distributions for each gestational age, sex and parity group that are optimal with regard to mortality risk in the sense that the mode of the curve coincides with the birth weight at which the risk of mortality is minimal.

2.1.2 Theoretical framework

In a study on good practice in addressing the problem of low birth weight,

Richardson *et al.* (2014) stated the reduction in child mortality is necessary in order to attain sustainable development goals. They identified the existence of a major challenge in the procurement of healthcare services by individuals which is determined to a large extent by their level of income. In their study, infant mortality rate, under- five mortality rate and neonatal mortality rate were modeled against household income and controlled for access to anti-natal care, access to safe water and sanitation, neonatal mortality rate, maternal education and household size in Nigeria. The findings of their study revealed that household income has significant effect on neonatal mortality rate in Nigeria but household income has insignificant effect on infant and under-five mortality rates in Nigeria. Also, it was found that household size has significant effect on infant mortality rate and neonatal mortality rate in Nigeria In addition, findings revealed that access to anti-natal care has significant effect on under-five mortality rate in Nigeria.

Amzat and Adeosun (2014) examined the nature of relationship between infant mortality and some socioeconomic and demographic variables. Also, assessed the proximate covariate that influences the survival of an infant using the 2003 Nigeria Demographic and Health Survey Data (NDHS). They used sequential probity model to examine the relationship between the dependent variables (infant's death and age at death) and predictor variables for both correlated and uncorrelated error terms. The findings of their study showed that in both of the situation with correlated and uncorrelated error terms, infant's being alive or death is positively affected by education, birth order number, duration of breast feeding and negatively affected by both total children born and place of delivery. There exist significant differences among the predictor variables on the probability of infant's death at neonatal and post neonatal period. Also, the correlation between the error terms was found to be significant.

Adetoro and Amoo (2014) stated that despite the global decline in infant mortality rate from 90 deaths per 1,000 live births in 1990 to 48 in 2012, Nigeria is yet to record any substantial improvement. Infant mortality in Nigeria increased from 138 per 1,000 live births in 2007 to 158 per 1,000 live births in 2011 against the Millennium development Goal target of 71 per 1,000 live births. They used data from the Nigeria Demographic and Health Survey (NDHS) 2008 to investigate the predictors of child (aged 0-4 years) mortality in Nigeria. They statistical tool employed were cross-tabulation and binary logistic regression techniques. The findings of their study showed that mortality rate was highest (49.14 percent) for children of illiterate mothers and lowest (13.29 percent) among mothers with higher education. Also, the result of the logistic regression analysis revealed that, education of both parents and occupation of mothers were found statistically significant to reduction in child mortality rate. It was equally found that mothers' wealth index, age at first birth and usual of place of residence have substantial impact on child mortality in Nigeria. They concluded that increase in women education could increase age at first birth and mitigate the risk of poor child health outcomes.

Adepoju (2015) examined the differentials in child mortality rate across socioeconomic, demographic and selected health characteristics in rural Nigeria, employing the 2008 National Demographic and Health Survey data. The findings of his study on health attributes and morbidity pattern of mother and child revealed that most of the respondents did not have access to good health facilities and antenatal care. As a result, more than three-quarters of the respondents delivered their babies at home and had less than 24 months birth interval between pregnancies. Results showed that child mortality rate was highest among illiterate mothers, mothers without a source of income, under aged women (less than 20 years) and among fathers whose primary livelihood lie in agriculture. Regional analysis showed that the North-Western zone had the highest child mortality rate followed by the North-Eastern zone, while the South-South zone had the lowest. With respect to health attributes, children delivered at home, who were never breastfed and of multiple births had high mortality rates. Gender differentials showed that the rate of mortality was higher for male than for female children but lowest for children who had been fully immunized and whose mothers were aged between 21 and 30 years.

Jacdonmi *et al*. (2016) reviewed the trends and patterns of breastfeeding, causes of infant mortality and breastfeeding of infants from birth to six months, followed by appropriate and adequate complementary feeding for two years and above, as a strategic intervention against infant mortality and the need to create awareness about the benefits of breastfeeding. The outcome of their review showed that breastfeeding protects infants from several infections such as diarrhea, pneumonia, gastrointestinal infections, urinary tract infections, sudden infant death syndrome and others which are probable causes of infant deaths. They noted that as breastfeeding provides adequate nutrition to infants, protects them from diseases and infections, it is a cost-effective method/intervention to reduce infant mortality.

Liu *et al.* (2016) assessed the extent and correlates of stillbirths being misclassified as neonatal deaths by Comparing two recent and linked population surveys conducted in Malawi, one being a full birth history (FBH) survey, and the other a follow-up verbal/social autopsy (VASA) survey. The result of their study found that one-fifth of 365 neonatal deaths identified in the FBH survey were classified as stillbirths in the VASA survey. Neonatal deaths with signs of movements in the last few days before delivery reported were less likely to be misclassified stillbirths ($OR = 0.08$, $p<0.05$). It was found that having signs of birth injury has impact on higher odds of misclassification ($OR = 6.17$, $p < 0.05$).

Yakubu *et al,* (2019) Binary Logistic Regression Methods for Modeling Broncho-Pneumonia Status in Infants from Tertiary Health Institutions in North Central Nigeria. Acute respiratory tract infections, predominantly bronchopneumonia, are one of the leading causes of infant deaths in developing countries and around the world. This work models the effects of the significant risk factors on infants' bronchopneumonia status and also fits some reduced models and determines the best model with minimum number of parameters. The data for this study consist of a random sample of 433 births to women seen in the obstetrics clinic of two sampled tertiary health institutions in north-central Nigeria. These include University Teaching Hospital (UTH) Abuja, and Federal Medical Center (FMC) Keffi, Nasarawa State. Binary logistic regression was used to identify and model the effects of the various risk factors while stepwise regression technique was used to fit some reduced logistic regression models. Then the best fitting model with minimum number of parameters was identified using likelihood ratio statistic. It was observed that baby's weight at birth, baby's weight four weeks since birth, and mother's occupation have significant effects on infant's bronchopneumonia status. Additionally, among the four fitted reduced models, model 4 is the best predictor of infants' bronchopneumonia status, followed by model 3 and then model 2. Therefore, community service like home visiting for health education, supplementation of vitamin A, etc., would be an advantage if provided for teenaged pregnant women as it would, in turn, reduce incidence of low birth weight and thereby reduce bronchopneumonia infection among these children.

Adeyemi *et al.* (2016) the study investigates the social and demographic (environmental, maternal and child characteristics impacts on the child birth weight. The approach also assesses the geographical variations in child birth weight across the states and the woman likelihood of having a child birth weight falling in a specific category. The fixed effect estimates of regression models look reasonable and controlled for all stable characteristics of the mother including household wealth index and child-specific factors. The descriptive analysis showed that low birth weight delivery decreases with an increase in birth orders of second, third and fourth as compared to first order births, but not to the fifth order (5 or more). The incidence of low birth weight are inversely associated the household wealth index. The bivariate analysis showed that a large proportion of women did not attend antenatal and they did not give postnatal vitamin A supplements to their children. From the binary logit analysis, the findings revealed that mother literate and prenatal iron syrup supplementation had significant association with a lower probability for a low birth weight. Other variables include urban residence and antenatal attendance also had a strong influence with a low chance of low birth weight, but they were not significant in the study.

The findings of the study also revealed that the childhood under nutrition and disease had significant association with higher likelihood for the low birth weight. Evidence had showed that child birth defects may be genetically induced or environmental inherent. A recent study had reported that an underweight mother had a higher risk of giving birth to an underweight baby.by the result corroborates.

The study also revealed that firewood/dung cooking and mother smoking (tobacco) are critical risk factors of low birth weight in Nigeria. This result is complementary to early work conducted in Zimbabwe. They asserted that low birth weight not only caused by lack of socioeconomic resources but by the use of inferior energy sources for indoor cooking and air pollution Other previous studies had enunciated that early childbearing, inadequate access to prenatal health services and less disadvantage groups experienced a higher prevalence of low birth weight in cities than in rural areas in sub-Saharan Africa. The result gave a strong indication that iron syrup supplementation during pregnancy yields improvement in child birth weight, while postnatal vitamin A intervention would boost the child growth of those children born with low birth weight. There was a wealthy of evidence that zinc supplementation reduces diarrhea morbidity and respiratory infections among children.

Kowsher *et al.* (2021) A woman's satisfaction with childbirth may have immediate and long-term effects on her health as well as on the relationship with her newborn child. The mode of baby delivery is genuinely vital to a delivery patient and her infant child. It might be a crucial factor for ensuring the safety of both the mother and the child. During the baby delivery, decision-making within a short time becomes very challenging for the physician. Besides, humans may make wrong decisions selecting the appropriate delivery mode of childbirth. A wrong decision increases the mother's life risk and can also be harmful to the newborn baby's health. Computer-aided decision-making can be an excellent solution to this problem. The study has applied 32 supervised classifier algorithms and 11 training methods on the real childbirth dataset from the Tarail Upazilla Health complex, Kishorganj, Bangladesh. The study also analyzed the data and compared them using measure criteria to determine the best performed model. The analysis shown that quadratic discriminant had the highest accuracy of 0.97 with the F1 score of 0.97. Using this model to decide the appropriate labor mode may significantly reduce maternal and infant health risks

2.2 Gap identified in literature

The present work is different from other previous works as in Adler and Donlon (2020) in which a combination of discriminant and binary logistic regression were used and considered only three variables (sex, weight at birth and weight after four weeks) as against the present study in which seven variables were used. In Eneh and Ugwu (2021) the study investigated some variables considered from mothers' aspect and used logistic regression.

This study considered seven maternal characteristics to determine which variables have significant in classifying the response variable

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Study Design

To achieve the research objectives, this work is a combination, both in purpose and in design of discrimination and classification analysis. It is discrimination as it seeks to draw a line between the birth weight status of infants using their maternal height, maternal weight, maternal age, baby's weight, baby's sex, gestational age and parity. On the other hand, in the classification design, the researcher is not interested in a mere collection of haphazard facts but model would be used to classify the birth weight status of an infant whose birth weight is not known earlier.

However, in discrimination and classification designs, the major statistical components form the basis of the research design which includes both the sampling plan and the modeling procedures. The sampling plan is the methodology used for selecting the sample from the population. The modeling procedure is the algorithms or formulae used for obtaining models of population values from the sample data and for estimating the reliability of these models.

The study considered seven selected predictor variables which are capable of characterizing birth weight in new babies. From experience and records of medical practice, these variables have shown significantly affect/discriminate between normal births (π_1) and low birth weight (π_2).

These variables included in the model are: X_{mh} = maternal height (cm), X_{mw} = maternal weight, X_{ma} = maternal age, X_{bw} = baby's weight, X_{bs} = baby's sex, X_{ga} = gestational age and parity

3.2 Population and Sample of Study

This study was conducted on the available data from the Jummai Babangida Aliyu Neonatal and Maternal Hospital Minna. The study is restricted to the following birth weight. Low birth weight (LBW) and normal birth weight (NBW). A total of 608 data files were available for inspection and a total sample of 218 was collected for the study

3.3 Data Collection

This research work was carried out in Minna, where data were carefully and technically extracted directly from the individual client's folder from Jummai Babangida Aliyu Neonatal and Maternal Hospital Minna. The maternal height, maternal weight, maternal age, baby's weight (kg), baby's sex, gestational age and parity. The data is of secondary source, the sampling design and data collected procedure adopted in this research work is the simple random sampling (SRS) scheme with size n = 2. In this sampling method, each member of the population has an exactly equal chance of being selected, the study adopt simple random number table to make sample of 218 from the available record of 608

3.4 Discriminant Analysis Procedure

Given a set of p independent variables $X_1, X_2,...,X_p$ (Maternal characteristics in this case), the technique attempts to derive a linear combination of these variables (maternal characteristics) which best separates or discriminates the two groups (low birth weight and normal birth weight). The functions are generated from a sample of cases for which group membership is known; the functions can then be applied to new cases with measurements for the predictor variables, but unknown group membership.

The Discriminant function for this study is expressed in the form

$$Z = W_0 + W_1X_1 + W_2X_2 + \cdots + W_kX_k \tag{3.1}$$

Where: Z = discriminant score; W_0 = discriminant constant; W_k= discriminant weight or coefficients; $X_k =$ an independent variable or predictive variables and k = objects in groups

The procedure automatically chooses a first function that will separate the groups as much as possible, it then chooses the second function that is both uncorrelated with the first function and provides as much further separation as possible. The procedure continues adding functions in this way until reaching the maximum number of functions as determined by the number of predictors and groups in the dependent variable. In two group discriminant function, there is only one discriminant function. The discriminant score obtained from the discriminant function shall classify the birth weight into one of the two groups.

Mbanasor *et al,* (2008) testing the classification performances of the discriminant function, use the overall hit ratio which is the same thing as percentage of the original group cases which correctly classified, the dependent variables using the standardized discriminant coefficients. The greater the magnitude of the coefficients, the greater the impact of the variable as an identifying variable. However, to test the significance of the discriminant function as a whole we shall use the Wilks' Lambda. The analysis of variance (ANOVA) table for the discriminant function score is another overall test of the discriminant analysis model. If the probability value (p-value) is less than 0.05 (level of significance)

Assumi ng there are two multivariate normal populations with equal variance-covariance matrices, $N(\mu_1, \sum_1)$ and $N(\mu_2, \sum_2)$ where $\mu_i (i = 1,2) = (\mu_1, \mu_2)$ is the vector of means of the i^{th} population and $\sum$ is the common variance-covariance matrices of the two populations. The probability density function of ith population is given as follow:

$$P_i(X) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\tfrac{1}{2}(X-\mu_i)' \Sigma^{-1} (X-\mu_i) \right] \tag{3.2}$$

The ratio of the densities of two multivariate normal populations is given below Usman (2011):

$$\frac{P_1(X)}{P_2(X)} = \frac{\exp\left[-\tfrac{1}{2}(X-\mu_1)' \Sigma^{-1} (X-\mu_1) \right]}{\exp\left[-\tfrac{1}{2}(X-\mu_2)' \Sigma^{-1} (X-\mu_2) \right]} \geq k$$

$$\exp\left[-\tfrac{1}{2}\left\{ (X-\mu_1)' \Sigma^{-1} (X-\mu_1) - (X-\mu_2)' \Sigma^{-1} (X-\mu_2) \right\} \right] \geq k \tag{3.3}$$

By taking the natural logarithms of equation (3.3) above; which is monotone increasing we have:

$$-\tfrac{1}{2}\left\{ (X-\mu_1)' \Sigma^{-1} (X-\mu_1) - (X-\mu_2)' \Sigma^{-1} (X-\mu_2) \right\} \geq \log k \tag{3.4}$$

The second term of (3.4) above is the Mahalonobis square distance between $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$. For k suitably chosen (which of course can be one and then log k will be zero), the left hand side of equation (3.4), can be expanded and repositioned to get the following equation:

$$X' \Sigma^{-1}(\mu_1 - \mu_2) - \tfrac{1}{2}(\mu_1 + \mu_2)' \Sigma^{-1}(\mu_1 - \mu_2) \geq \log k \tag{3.5}$$

The first expression of equation (3.5) above is the well known as Fisher's linear discriminant function which is linear in the component of the observation vector.

3.4.1 Method used in the study

In this work, seven predictor variables that are well recognized for characterizing birth weight infant were considered. These variables are maternal height, maternal weight, maternal age, baby's weight (kg), baby's sex, gestational age and parity By the method of Euclidean distance, the mean vectors and the covariance matrices of a sample of both low birth weight (π_1) and normal birth weight (π_2) as cited in Usman (2011);

Let

$$\overline{X}_i = \begin{pmatrix} \overline{x}_{i1} \\ \overline{x}_{i2} \\ ,, \\ ,, \\ \overline{x}_{ip} \end{pmatrix} \quad (3.6)$$

Where $\overline{X_i}$ represent the sample mean vector and i denote the two groups (LBW and NBW). Let $\overline{X}_{i1}, \overline{X}_{i2} \dots, \overline{X}_{ip}$ represent the individual mean vectors for the seven variables i.e. p = 7

For instance;

$$\overline{X}_{i1} = \frac{1}{k}\sum_{j=1}^{n} X_{i1} \quad (3.7)$$

where X_{11} is the mean of the first variable in group one, while X_{21} represent the mean of the first variable in group two, k is the number of the cases and n is the sum of all observations in particular group,

The sample variance-covariance matrix is given as;

$$\mathbf{S}ij = \begin{pmatrix} S_{11} & S_{12} & . & . & . & S_{1p} \\ S_{21} & S_{22} & . & . & . & S_{2p} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ S_{p1} & S_{p2} & . & . & . & S_{pp} \end{pmatrix}$$

(3.8)

where S_i denotes variance-covariance matrix, for $i = 1, 2$
S_{ii} denotes an individual variance and
S_{ip} denotes an individual covariance for $p = 1, 2, \ . \ . \ ., 7.$

The illustrations are given below,

$$S_{ij} = \frac{1}{k_i}\sum_{j=1}^{ni}\left(X_{ij} - \overline{X}_{.i}\right)^2 \quad \text{(general variance)} \tag{3.9}$$

$$S_{11} = \frac{1}{k}\sum_{j=1}^{n}\left(X_{i1} - \overline{X}_1\right)^2 \quad \text{(an individual variance)} \tag{3.10}$$

$$S_{12} = \frac{1}{k}\sum_{i=1}^{n}\left(X_{i1} - \overline{X}_1\right)\left(X_{i2} - \overline{X}_2\right) \text{ (an individual covariance)} \tag{3.11}$$

Let π_1 denotes group one (low birth weight) and

π_2 denotes group two (normal birth weight)

The Euclidean distance of the low birth weight (π_1) is;

$$\hat{l}_1 = X_1' S_p^{-1}(\overline{X}_1 - \overline{X}_2) \tag{3.12}$$

and Euclidean distance of the normal birth weight (π_2) is;

$$\hat{l}_2 = X_2' S_p^{-1}(\overline{X}_1 - \overline{X}_2) \tag{3.14}$$

Where S_p denotes the pooled variance matrix

The mean Euclidean distance used in this study for the two groups is given as;

$$\widehat{M} = \frac{1}{2}\left(\hat{l}_1 + \hat{l}_2\right) \tag{3.15}$$

and the Discriminant function is calculated by

$$\hat{Y} = X' S_p^{-1}(\overline{X}_1 - \overline{X}_2) \tag{3.16}$$

Therefore, the classification rule is that;

$if \quad \hat{Y} \geq \widehat{M} \quad classified\ as\ group\ one\ (\pi_1)$ and

$if \quad \hat{Y} < \widehat{M} \quad classified\ as\ group\ two\ (\pi_2)$

Where $\hat{Y}$ denote the Discriminant function, and $\widehat{M}$ denote the mean Euclidean distance for Normal birth and Low birth weight groups

$$\mathbf{X}' = (X_1 \quad X_2) \tag{3.17}$$

$$S_p = \frac{n_1 S_1 + n_2 S_2}{n_1 + n_2} \tag{3.18}$$

Since $n_1 \neq n_{2,,}$equation (3.18) will be used but if $n_1 = n_2$ then the estimated pooled variance S_p above becomes:

$$S_p = \frac{S_1 + S_2}{2} \tag{3.19}$$

Where $S_1 and\ S_2$ are the respective sample variance covariance matrices for the two groups, and $n_1 and n_2$ are the sample size of the two groups respectively.

The Fisher's linear Discriminant model used is:

$$y_{ij} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + e$$

(3.20)

$i = 1,2, \dots ,208,$

$j = 1,2, \ldots ,7$

where;

y_{ij} = denote response probability (birth weight), x_1 = maternal height (cm), x_2 = maternal weight (kg), x_3 = maternal age, x_4 = baby's weight, x_5 = baby's sex, x_6 = gestational age, x_7 = parity, β_0 = expected value of y when x's are set at zero, β_i = regression coefficient for each corresponding predictor variable and ε = error of the predictor

$$y_{ij} = \begin{cases} 1 \; the\ child\ weight\ < 2.5 \\ 0 \; the\ child\ weight\ > 2.5 \end{cases}$$

$$p(y=1/x) \tag{3.21}$$

3.4.2 Test for significance of canonical correlation (wilk's lambda)

The degree of linear relationship existing between two variables can be measured by means of Canonical Correlation which takes on values between minus-one and plus-one inclusive ($-1 \le r \le +1$). The closer the value of Canonical Correlation is to one, the stronger the degree linear relationship existing between the two set of variables. Also, the stronger the degree of linear relationship existing between the two set of variables, the better the linear discriminant function between the set of variables. The significance of Canonical Correlation is measured by the Wilk's' Lambda statistic as follows;

Hypothesis for Canonical Correlation:

H_0: There is no linear relationship between the two set of variables

H_1: There exists a linear relationship between the two set of variables

Test statistic:

$$\lambda = \frac{|W|}{|W + H|} \tag{3.22}$$

Where;

W is residual variance

H is variance due to linear relationship

W+H is the total variance.

Decision Rule:

Reject H_0 if $p < 0.05$ otherwise accept H_0 at the 5 percent level of significance.

In this research work, we investigated the strength of the underlying relationship between several pairs of variables, using canonical correlation coefficient whenever such relationships are to be measured.

3.4.3 Omnibus chi-square test

The omnibus Chi-square test is a log-likelihood ratio test for investigating the model coefficients in logistic regression. The test procedures are as follows:

Hypothesis for Omnibus Chi-square test:

H_0**:** The model coefficients are not statistically significant

H_1**:** The model coefficients are statistically significant

Test statistic:

$$\chi^2 = 2\left[\sum_{i=1}^{r}\sum_{j=1}^{c} O_{ij} \ln\left(\frac{O_{ij}}{e_{ij}}\right)\right] \tag{3.24}$$

Where;

O_{ij} denote observed values and e_{ij} denote expected values

O_{ij} *denote Observed Values, and* e_{ij} *denote Expected values*

Decision Rule:

Reject H_0 if $p < 0.05$ otherwise accept H_1 at the 5 percent level of significance.

3.4.4 Box's M - test for the equality of covariance matrices

Box's M test was used to determine whether two or more covariance matrices are equal. It is also used to test for homogeneity of covariance matrices. The basic assumptions of the linear discriminant model are that the two covariance matrices must be equal. Hence, the Box M test is used to investigate this assumption; otherwise the discriminant model may be misleading Jiamwattanapong *et al.* .(2021)

Hypothesis for Box's M Test:

H_0: The two covariance matrices are not equal

H_1: The two covariance matrices are equal

Test statistic:

$$M = \frac{|S_L|}{|S_S|} \tag{3.25}$$

where S_l and S_s are the larger and smaller variance respectively

Decision Rule:

Reject H_0 if $p < 0.05$ otherwise accept H_1 at the 5% level of significance.

$$P = \Pr\left(F_{v1,,v2} > M\right)$$

Where;

M = Calculated value of Box's M

$F_{v1,v2}$ = F-distribution with v_1,v_2 df

The Box's M was used to investigate the equality of the two covariance matrices. That is, if they are equal, then the linear discriminant model is appropriate otherwise the non-linear discriminant model is applied. Hence, the Omnibus test is applied.

3.4.5 Logistic regression model

Logistic regression or logit deals with the binary case, where the response variable consists of just two categorical values. Logistic regression model is mainly used to identify the relationship between two or more explanatory variables. X_i and the dependent variable Y. Logistic regression model has been used for prediction and determining the most influential explanatory variables on the dependent variable. The logistic regression model is the most frequently used regression model for the analysis of these data. It is important to understand that the goal of an analysis using this model is the same as that of any other regression model used in statistics, that is, to find the best fitting and most parsimonious, clinically interpretable model to describe the relationship between an outcome (dependent or response) variable and a set of independent (predictor or explanatory) variables.

The most important difference between a logistic regression model and the linear regression model is that the outcome variable in logistic regression model is *binary* or *dichotomous w*hile, in linear regression model it is assumed that an observation of the outcome variance may be expressed as $E(\varepsilon) = 0$. This difference between logistic and linear regression is reflected both in the form of the model and its assumptions.

Unlike linear regression, which predicts the actual values of the response variable, logistic regression models the probability associated with each level of the response variable by finding a linear relationship between predictor variables and a link function of these probabilities. Different link functions offer different goodness-of-fit for the data. The following link functions are common, and during data analysis, the link function that offers the best goodness-of-fit for the data is chosen.

Binary and ordinal logistic regression offers all three link functions; nominal logistic regression offers only the logit link function. In order to simplify notation, we use the quantity $\pi(x) = E(Y/x)$ to represent the conditional mean of Y given x when the logistic distribution is used.

Model fitting: if Y denotes baby's weight at birth with values "1" if the baby has Normal birth weight (*a success*) and "0" otherwise Low birth weight (*a failure*), for every sampled infant, the probability that he/she has normal birth weight (i.e., *a success*) is $\pi(\mathrm{x}) = \mathrm{P}(\mathrm{y} = 1/\mathrm{X})$ and the corresponding probability that he/she has low birth weight (*a failure*) is $1 - \pi(\mathrm{x}) = \mathrm{P}(\mathrm{y} = 0/\mathrm{X})$ The logistic regression model used is:

$$\pi(\mathrm{x}) = \frac{e^{\beta_0+\beta_1 \mathrm{x}_i}}{1 + e^{\beta_0+\beta_1 \mathrm{x}_i}}$$

$$i = 1,2,3, \dots ,7$$

$$\pi(\mathrm{x}) = \frac{e^{\beta_0+\beta_1 X_{mh}+\beta_2 X_{mw}+\beta_3 X_{ma}+\beta_4 X_{bw}+\beta_5 X_{bs}+\beta_6 X_{ga}+\beta_7 X_p}}{1+e^{\beta_0+\beta_1 X_{mh}+\beta_2 X_{mw}+\beta_3 X_{ma}+\beta_4 X_{bw}+\beta_5 X_{bs}+\beta_6 X_{ga}+\beta_7 X_p}} \tag{3.26}$$

Where $\hat{\pi}(x_i)$ is the predicted probability of the i^{th} infant at x_i; X_{mh}, X_{mw}, X_{ma}, $X_{b'w}$, $X_{b's}$, X_{ga} and X_p denote respectively, maternal height, maternal weight, maternal age, baby's weight, baby's sex, gestational age and parity. $\hat{\beta}_0$ denotes the estimated intercept and β_h, $h = 1,2,3, \dots p$ denotes the logistic regression coefficient for the i^{th} predictor variables.

Since model (3.26) is nonlinear, the logit transformation on

$\acute{\pi}(x_i)$ yields the multiple logistic regression model as follows

$$\dot{g}(\mathrm{x}) = logit\left(\acute{\pi}(x_i)\right) = \ln\left[\frac{\pi(x_i)}{1 - \pi(x_i)}\right]$$

$$= \hat{\beta}_0 + \sum_{i=1}^{7} \hat{\beta}_1 x_i$$

$i = 1,2,3, \dots ,7$

The importance of this transformation is that *g(x)* has many of the desirable properties of a linear regression model. The logit, *g(x)*, is linear in its parameters, may be continuous, and may range from $-\infty$ to $+\infty$, depending on the range of *x*.

Logistic regression models are adequate for those situations where the dependent variable of the regression problem is binary. That is, the dependent variable has only two possible outcomes, e.g "success/failure" or "normal/abnormal". We assumed that these binary outcomes are coded as 1 and 0

The application of linear regression models to such problems would not be satisfactory since the fitted predicted response would ignore the restriction of binary taking on values for the observed data. In this work, an attempt is made to estimate a population regression equation as;

$$y_{ij} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \varepsilon \qquad (3.27)$$

The response Y_{ij} is continuous, and is assumed to follow a normal distribution. The study will predict the mean value of the response corresponding to a given set of values for the explanatory variables.

However, there are many situations in which the response of interest is dichotomous rather than continuous. Examples of variables that assume only two possible values are disease status (the disease is either present or absent) and survival following surgery (a patient is either alive or dead).

Our interest is to estimate the probability (p) associated with a dichotomous response (which, of course, is also its mean) for various values of an explanatory variable.

In this situation, the study only considered simple logistic regression– that is, logistic regression models with explanatory variables. The first strategy might be to fit a model of the form:

$$P_{BW} = \beta_0 + \beta X_i + \varepsilon \tag{3.28}$$

$$where\ Xi = (MH,\ MW,\ MA,\ B'sW,\ B'sS,\ GA\ and\ P)$$

This is simply the standard linear regression model in which X_i represent the explanatory variables and Y the outcome of a continuous normally distributed random variable. Where α is the intercept and β is the slope. On inspection, however, this model is not feasible. Since p is a probability, it is restricted to taking values between '0' and '1'.

3.5 Wald's Test

The Wald's test is a parametric statistical test named after the Transylvanian statistician Wald (1943) with a great variety of uses. Whenever a relationship within or between data items can be expressed as a statistical model with parameters to be estimated from a sample, the Wald test can be used to test the true value of the parameter based on the sample estimate. Under the Wald statistical test, the maximum likelihood estimate of the parameter(s) of interest is compared with the proposed value, a logistic regression model. Deviance is calculated by comparing a given model with the saturated model a model with a theoretically perfect fit. This computation is called the likelihood-ratio test assumption that the difference between the two will be approximately normally distributed. Typically the square of the difference is compared to a chi-squared distribution.

The Wald statistic is given as

$$\boldsymbol{w_j = \frac{B_J^2}{SeB_J^2}} \tag{3.29}$$

where $B_J^2$is the square of the slope and SeB_J^2 is the standard error of the slope

3.6 Hosmer- Lemeshow's Goodness of Fit

The Hosmer–Lemeshow test is a statistical test for goodness of fit for logistic regression models. It is used frequently in risk prediction models. The test assesses whether or not the observed event rates match expected event rates in subgroups of the model population. The Hosmer–Lemeshow test specifically identifies subgroups as the deciles of fitted risk values. Models for which expected and observed event rates in subgroups are similar are called well calibrated.

The Hosmer –Lemeshow test statistic is given by:

$$\boldsymbol{H} = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{E_{g\ (1-E_{g\ /Ng})}} \tag{3.30}$$

Here O_g, E_g, N_g, and π_g denote the observed events, expected events, observations, predicted risk for the g^{th} risk decile group, and G is the number of groups. The test statistic asymptotically follows a χ^2 distribution with $G - 2$ degrees of freedom. The number of risk groups may be adjusted depending on how many fitted risks are determined by the model. This helps to avoid singular decile groups. The test provides subjects into deciles based on predicted probabilities, and then compute a chi- square from observed and expected frequencies. Then a probability (P) value is computed from the chi- squared distribution. If the H-L goodness of fit test statistic is greater than 0.05, as we want for a well-fitting model we do not reject the null hypothesis that there is no difference between observed and model predicted values, implying that the models estimate fit the data at an acceptable level.

3.7 Confusion Matrix

Actual values

Predicted values

	Positive (1)	**Negative (0)**
positive(1)	**TP**	**FP**
Negative (0)	**FN**	**TN**

A confusion matrix is a table that is used to define the performance of a classification algorithm, its visualizes and summarizes the performance of classification algorithm

The following are the basic terminology which will help us in determining the metrics we are looking for

- True Positive (TP): When the actual value is positive and predicted is also positive
- True Negative (TN): When the actual value is negative and prediction is also negative
- False Positive (FP): When the actual value is negative but prediction is Positive. Also known as Type I error
- False Negative (FN): When the actual value is negative but prediction is negative. Also known as Type II error

3.8 Multicollinearity

A critical condition for the application of least squares is that the explanatory variables are not perfectly linearly correlated (i.e., $r_{xixj} \neq 1$). The term multicollinearity is used to denote the presence of linear relationship (or near linear relationships) among explanatory variables. If the explanatory variables are perfectly linearly correlated, that is, if the correlation coefficient for these variables is equal to unity, the parameters become indeterminate: it is impossible to separately obtain numerical values for each parameter and the method of least squares breaks down.

3.8.1 Effects/consequences of multicollinearity

Since $\hat{\beta} = (x^I x)^{-I} x^I y$ (3.30)

Where $(x^I x)^{-I} = \frac{Cof\ (x^I x)^T}{\det(x^I x)}$ (3.31)

If the $x's$ are highly correlated, then

i. $\det(x^I x) \rightarrow 0$

$=> (x^I x)^{-I} \rightarrow \infty$

$therrefore\ \beta \rightarrow \infty$

ii. $\text{var}(\beta) = s^2 (x^I x)^{-I} \rightarrow \infty$ (3.32)

Variance is infinite. This results in insignificant t-ratios

$$t^* = \frac{\hat{\beta}-\beta}{SE(\hat{\beta})} = \frac{\hat{\beta}-\beta}{\sqrt{\hat{\sigma}^2}} \quad (3.33)$$

iii. The variables of the parameter estimate are unnecessarily high.

CHAPTER FOUR

4.0 RESULTS AND DISCUSSIONS

4.1 Introduction

In this chapter, the data are fit to the linear Discriminant and logistic regression models. The results of the analyses are presented and discussed. The data were analyzed using SPSS version 28.0

4.2 Presentation and Discussion of Results

Table 4.1: Descriptive Statistics

	N	Mean	Std. Deviation
MATERNAL HEIGHT	608	1.5942	0.09807
MATERNAL WEIGHT	608	59.2572	9.16302
MATERNAL AGE	608	27.6694	4.99630
BABY’S WEIGHT	608	2.7495	0.62880
BABY’S SEX	608	1.4700	0.50000
GESTATIONAL AGE	608	34.1382	2.71380
PARITY	608	2.6800	2.14600

From Table 4.1, it is observed that the predictors with larger means are also associated with larger standard deviation. For instance, maternal weight has a larger mean of 59.2572 with corresponding standard deviation of 9.16302 and also the maternal age has a mean of 27.6694 and standard deviation of 4.99630 and so on. The tables are used to describe the basic features of the dataset, it provides simple summaries about the sample

Table 4.2: Correlational Matrices of the indicator variables (Pooled Within-Groups)

	MH	MW	MA	BW	BS	GA	P

MH	1.000	.493	.001	.053	-.050	-.004	-.017
MW	.493	1.000	.058	.141	-.041	.012	-.093
MA	.001	.058	1.000	-.015	-.037	-.023	.030
BW	.053	.141	-.015	1.000	.037	-.047	.057
BS	-.050	-.041	-.037	.037	1.000	-.026	-.063
GA	-.004	.012	-.023	-.047	-.026	1.000	-.096
P	-.017	-.093	.030	.057	-.063	-.096	1.000

From Table 4.2 indicate that, the data set meets up with the assumptions of Discriminant analysis which states that the predictors are not correlated with one another, that is, maternal weight, the correlation is 0.493 which is weak correlation and also the correlation between the self-predictor is constant across group.

Table 4.3: Pooled Covariance Matrices

Covariance	MH	MW	MA	BW	BS	GA	P
MH	.010	.437	.001	.002	-.002	-.001	-.003
MG	.437	82.393	2.638	.527	-.184	.286	-1.817
MA	.001	2.638	24.780	-.030	-.092	-.315	.317
BW	.002	.527	-.030	.169	.008	-.052	.051
BS	-.002	-.184	-.092	.008	.250	-.036	-.068
GA	-.001	.286	-.315	-.052	-.036	7.389	-.559
P	-.003	-1.817	.317	.051	-.068	-.559	4.604

Table 4.3 indicates that all within variables are significant i.e. MH to MW is 0.010 < .437. It also shows that, variable with other variables have at most one significance while the remaining are insignificant

Table 4.4: Test of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
MATERNAL HEIGHT	.994	3.821	1	605	.051
MATERNAL WEIGHT	.981	11.481	1	605	.001
MATERNAL AGE	.991	5.608	1	605	.018
BABYS WEIGHT	.426	814.856	1	605	.000
BABYS SEX	1.000	.053	1	605	.818
GESTATIONAL AGE	1.000	.008	1	605	.930
PARITY	.999	.738	1	605	.391

Table 4.4 measures the potential of variables, as small value indicate the variable that is better at discriminating between groups. It is observed that baby's weight at birth is best discriminating between the two groups since it has a smaller value of 0.426.

Table 4.5: Test Results of Box's M

Box's M		165.859
F	Approx.	5.840
	df1	28
	df2	644248.152
	Sig.	0.000

Table 4.5 as stated in equation 3.25, investigates the equality of the two covariance matrices. The test statistic is clearly stated in equation 3.25. The f-value of 5.840 and p-value of the Box's M of 0.000 has confirmed the equality of the covariance matrices for the two groups.

Wilk's lambda is a measure of how well each function separates cases into groups, it is equal to the proportion of total variance in the Discriminant scores not explained by differences among the groups. Smaller value of wilk's lambda indicates greater discriminating ability of the function.

Table 4.6: Wilk's Lambda Test

Function(s)	Test of Wilk's Lambda	Chi-square	Df	Sig.
1	.422	518.610	7	.000

Table 4.6 justifies the significance of the canonical correlation Wilk's Lambda statistic which gives 0.422 with p-value of 0.000. Comparing the p-value of Wilk's Lambda of 0.000 with the predefined significance level of $\alpha = 0.05$. Its measures how well each function separates cases into groups. It is equal to the proportion of the total variance in the discriminant scores not explained by differences among the groups. Smaller values of wilks' lambda indicate greater discriminatory ability of the functions

4.3 Linear Discriminant Function

The coefficients in the table below will be utilized to build up the model for the low birth weight (LBW) and Normal birth weight (NBW) group respectively.

Table 4.7: Fisher's Classification Function Coefficients

	BIRTH WEIGHT	
	LBW	NBW
MATERNAL HEIGHT	182.282	183.272
MATERNAL WEIGHT	0.362	0.376
MATERNAL AGE	1.219	1.269
BABY'S WEIGHT	12.219	18.231
BAB'YS SEX	8.453	8.321
GESTATIONAL AGE	4.920	4.965
PARITY	1.059	1.022
Constant	-255.341	-274.240

Table 4.7 provides information (coefficients) that will be used to create model for the low birth weight and normal birth weight groups. It shows that maternal height has the highest value (182.282 and 183.272) in both low birth weight and normal birth weight respectively while, parity has the least coefficients. In normal birth weight group, the response to birth weight is positive when other variables stand as zero (0). The positive constant value in low birth weight group indicate the increases in response (y) as predictors (x) increases**.**

The Fisher's linear discriminant model for each group is constructed as follows;

low birth weight (π_1)

$$y_1 = -255.341 + 182.282X_{mh} + 0.362X_{mw} + 1.219X_{ma} + 12.219X_{b'sw} + 8.453X_{b'ss} + 4.920X_{ga} + 1.059X_p \quad (4.1)$$

Model 4.1 interpretations

- with the increase in maternal height by 1metre, it is expected to have about 182.282 increase in low birth weight when other factors are held constant
- with increase in maternal weight by 1kg, it is expected to have about 0.362 increase in low birth weight when other factors are held constant
- with increase in in maternal age by 1 year, it is expected to have about 1.219 increase in low birth weight when other factors are held constant
- with increase in gestational age by 1 week, it is expected to have 4.920 increase in low birth weight when other factors are held constant
- with increase in parity by 1, it is expected to have about 1.059 increase in low birth weight when other factors are held constant

$normal\ birth\ weight\ (\pi_2)$

$$y_2 = -274.240 + 183.272X_{mh} + 0.376X_{mw} + 1.269X_{ma} + 18.231X_{b'sw} + 8.321X_{b'ss} + 4.965X_{ga} + 1.022X_p \quad (4.2)$$

Model 4.2 interpretations

- with the increase in maternal height by 1metre, it is expected to have about 183.272 increase in normal birth weight when other factors are held constant
- with increase in maternal weight by 1kg, it is expected to have about 0.376 decrease in normal birth weight when other factors are held constant
- with increase in in maternal age by 1 year, it is expected to have about 1.269 increase in normal birth weight when other factors are held constant
- with increase in gestational age by 1 week, it is expected to have 4.965 increase in normal birth weight when other factors are held constant

- with increase in parity by 1, it is expected to have about 1.022 increase in normal birth weight when other factors are held constant

Table 4.8: Standardized Canonical Discriminant Function Coefficients

	Function 1
MATERNAL HEIGHT	0.039
MATERNAL WEIGHT	0.054
MATERNAL AGE	0.101
BABY'S WEIGHT	1.004
BABY'S SEX	-0.027
GESTATIONAL AGE	0.049
PARITY	-0.032
Constant	-6.280

From Table 4.8 Standardized coefficients allows you to compare variables measured on different scales. Coefficient with larger absolute values corresponds to variable with greater discriminating ability therefore Table 4.8 suggests that birth weight has a large coefficient value of 1.004, so it has a greater discriminating ability.

The coefficients in table 4.8 are used to generate model 4.3;

$$Y_{BW} = -6.280 + 0.039X_{mh} + \ 0.054X_{mw} + \ 0.101X_{ma} + \ 1.004X_{b'sw} \ - 0.027X_{ss} + 0\ .049X_{ga} \ - 0.032X_p \qquad (4.3$$

Model 4.3 interpretations

- with the increase in maternal height by 1metre, it is expected to have about 0.039 increase in baby's weight at birth when other factors are held constant

- with increase in maternal weight by 1kg, it is expected to have about 0.054 decrease in baby's weight at birth when other factors are held constant
- with increase in maternal age by 1 year, it is expected to have about 0.101 increase in baby's weight at birth when other factors are held constant
- with increase in gestational age by 1 week, it is expected to have 0 .049 increase in baby's weight at birth when other factors are held constant
- with increase in parity by 1, it is expected to have about 1.022 increase in baby's weight at birth when other factors are held constant

Table 4.9: Functions at Group Centroids

Birth weight	Function 1
LBW	-1.617
NBW	.843

Unstandardized canonical discriminant functions evaluated at group means. These are the means of the discriminant function scores by each group.

The Cutoff point ($\hat{\mathbf{M}}$) is computed as follows;

$$\hat{\mathbf{M}} = \tfrac{1}{2}\left(\overline{Y}_1 + \overline{Y}_2\right) = \tfrac{1}{2}\left(-1.617 + 0.843\right) = -\mathbf{0.}3869 \tag{4.4}$$

Therefore, the classification rule is stated as;

Classify as group 1 (Low birth weight) if $Y_{BW} \geq -0.3869$

Classify as group 2 (Normal birth weight) if $Y_{BW} < -0.3869$

Table 4.10: Prior Probabilities for Groups

Birth weight	Prior probabilities	Cases Used in Analysis
LBW	0.500	208
NBW	0.500	399
Total	1.000	607

From Table 4.10 indicates the prior probability of misclassifying birth weight to NBW at $\alpha = 0.05$ and prior probability of misclassifying LBW is also at $\alpha = 0.05$

Table 4.11: Classification Results

			Predicted Group Membership		
		Birth weight	LBW	NBW	Total
Original	Count	LBW	168(80.7%)	40(19.2%)	208(100%)
		NBW	59(14.9%)	337(85.1%)	396(100%)

From Table 4.11 the classification shows the practical result of using Discriminant model of the cases used to create the model. In this research work 608 mothers of the dataset was used in creating the model. From the selected cases 168 of 208 i.e. 80.7 percent of Low birth weight were correctly classified and 40 19.2 percent) were misclassified. 337 of 396 i.e. 85.1 percent of Normal birth weight were correctly classified while 59 i.e. 14.9 percent were misclassified.

4.4 The constructed logistic regression model

Logistic regression deals with the binary cases, where the response variable consists of just two categorical values. Logistic regression model is mainly used to identify the relationship between two or more explanatory variables. X_i and the dependent variable Y.

Table 4.12: Variables in the Equation for the Sample Data

	B	S.E.	Wald	df	Sig.	Exp (B)
MH	0.066	4.800	0.000	1	0.989	1.068
MW	0.011	0.053	0.047	1	0.008	1.011
MA	-0.001	0.076	0.000	1	0.991	0.999
BW	33.870	5.260	41.465	1	0.000	5.60
BS(1)	-0.029	0.720	0.002	1	0.968	0.972
GA	0.035	0.131	0.073	1	0.027	1.036
PA	-0.188	0.166	1.277	1	0.258	0.829
Constant	-83.142	15.769	27.800	1	0.000	0.000

Table **4.12** indicates parameter estimate of the model. Using the Wald's statistic, two coefficients are statistically significant while, the remaining are insignificant at $\alpha = 0.05$. From equation 3.27, we obtain the logistic regression model as follows: Model fitting: if Y denotes baby's weight at birth with values "1" if the baby has Normal birth Weight (*a success*) and "0" otherwise Low Birth Weight (*a failure*), the for every sampled infant, the probability that he/she has Normal Birth Weight (i.e., *a success*) is $\pi(x) = P(Y = 1/x)$ and the corresponding probability that he/she has Low Birth Weight (*a failure*) is $1 - \pi(x) = P(Y = 0/x)$.

$$\hat{\pi}(x_i) = \frac{e^{-83.142+.066X_{mh}+.011X_{mw}-.001X_{ma}+33.870X_{b\prime w}-.029X_{b\prime s}+.035X_{ga}-.188X_{p}}}{1+e^{-83.142+.066X_{mh}+.011X_{mw}-.001X_{ma}+33.870X_{b\prime w}-.029X_{b\prime s}+.035X_{ga}-.188X_{p}}} \qquad (4.5)$$

Where $\hat{\pi}(x_i)$ is the predicted probability for the i^{th} infant at X_i; $X_{mh}, X_{mw}, X_{ma}, X_{b\prime w}, X_{b\prime s}, X_{ga}, and\ X_p$ denote, respectively, maternal height, maternal weight, maternal age, baby's weight, baby's sex, gestational age and parity. $\hat{\beta}_0$ denotes the estimated intercept and β_h, $h = 1,2,3, \ldots p$ denotes the logistic regression coefficient for the i^{th} predictor variables.

Since model (4.5) is nonlinear, the logit transformation on $\hat{\pi}(x_i)$ yields the multiple logistic regression model as follows

$$\hat{g}(x) = logit\left(\hat{\pi}(x_i)\right) = \ln\left[\frac{\pi(x_i)}{1-\pi(x_i)}\right]$$

$$= \hat{\beta}_0 + \hat{\beta}_1 X_{mh}, + \hat{\beta}_2 X_{mw} + \hat{\beta}_3 X_{ma} + \hat{\beta}_4 X_{b'w} + \hat{\beta}_5 X_{b's} + \hat{\beta}_6 X_{ga} + \hat{\beta}_7 X_p$$

Where,
$\hat{\beta}_0 = -83.142$, $\hat{\beta}_1 = 0.066$, $\hat{\beta}_2 = 0.011$, $\hat{\beta}_3 = -0.001$, $\hat{\beta}_4 = 33.870$, $\hat{\beta}_5 = -0.029$, $\hat{\beta}_6 = 0.035$ and $\hat{\beta}_7 = -0.188$ (4.6)

Model 4.6 Interpretations

- with the increase in maternal height by 1metre, it is expected to have about 0.066 increase in baby's weight at birth when other factors are held constant
- with increase in maternal weight by 1kg, it is expected to have about 0.011 increase in baby's weight at birth when other factors are held constant
- with increase in in maternal age by 1 year, it is expected to have about -0.001 decrease in baby's weight at birth when other factors are held constant
- with increase in gestational age by 1 week, it is expected to have 0.035 increase in baby's weight at birth when other factors are held constant
- with increase in parity by 1, it is expected to have about -0.188 decrese in baby's weight at birth when other factors are held constant

Table 4.13: Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	10.848	8	.233

This Table **4.13** above suggests that the model is a good fit to the data since p = 0.233 > 0.05. However, the chi-square statistic shows that the weight depend on the categorical data

4.5 Checking for Multicollinearity

To check for presence of multicollinearity in the in dependable variables; the study correlate these variables with one another and obtained the results below;

Table 4.14 **Testing for Multicollinearity**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.852	4.253		.671	.503		
	MH	-.013	.018	-.065	-.714	.476	609	1.643
	MW	-.031	.142	-.020	-.218	.828	.621	1.610
	MA	-.084	.345	-.052	-.245	.807	.109 9.	155
	B'sW	-.113	.324	-.075	-.351	.726	.110	9.055
	B'sS	.088	.043	.232	2.051	.042	390	2.563
	GA	.434	.342	.231	3.453	.223	.221	4.554
	PA	-.212	-123	. 054	4.443	.432	.432	3.112

From the Table 4.14 above based on the coefficients output collinearity statistics, obtained VIF value of 1.643 and 1.610 respectively because the VIF value obtained is between 1-10, it can be concluded that there is no symptoms of multicollinearity

4.6 Major Findings

The following are the summary of findings;

1. No perfect correlation among the independent variables (i.e. no presence of multicollinearity).
2. From the selected cases 168 0f 208 i,e 80.7 percent of the low birth weight were correctly classified and 40 i.e 19.2 were misclassified. 337 of 396 i,e 85.1 percent of normal birth weight were correctly classified while 59 i.e 14.9 percent were misclassified.
3. Parameter estimates of the model using Wald's statistic, two coefficients are statistically significant while, the remaining are insignificant at $\alpha = 0.05$

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

In this research, Linear Discriminant and Logistic Regression Model were applied to data collected for birth weight from Minna, Niger State. The result shows that the prediction of birth weight is better done with Discriminant model than Logistic regression method

In this research, it was found that Discriminant model has a perfect classification of new cases than Logistic regression model. While, the reviewed models when tested with new cases, observed that, the Dicsriminant model has a perfect classification than the Logistic regression model.

5.2 Recommendations

The study recommend that the models developed in this study could assist the Doctors and other health practitioners to detect and monitor the prevalence and control of birth weight among infants

It is also recommended for further research and use of other statistical package especially those dedicated to multivariate analysis on this area in order to elucidate intensive information or results.

The recommendation also stated that Doctors and Clinics should adopt the use of the models built by this study to discover the prevalence of low birth weight among infants so that adequate measures for prevention and control of birth weight can be taken early enough

5.4 Contribution to knowledge

This study investigated the effect of seven different maternal characteristics variables (five continuous and two categorical) on the birth weight status of infants. The characteristics variables are maternal height, maternal weight, maternal age, birth weight, baby's sex, gestational age and parity. The maternal age, gestational age and parity were considered as new variables. The study is to discover the variable that is best discriminating between the two groups. As stated earlier, the data for the study were collected from Jummai Babangida Maternal and Neonatal Hospital Minna, Niger State Nigeria using simple random sampling scheme. The Discriminant and logistic regression model were used for the study.

The study extended the maternal characteristics variables to seven as against variables considered from other literatures where two and three maternal characteristics were considered

REFERENCES

Abalkhail, A. & Khalid, M. (1995). Low Birth Weight in the Taif Region: *Eastern Mediterranean Health Journal*, 1(1), 47–54.

Adepoju, A. O. (2015). Differential Pattern in Child Mortality Rate in Rural Nigeria. *Annual Research and Review in Biology, 7(5), 309-317.*

Adetoro, G. W. & Amoo, E. O. (2014). A Statistical Analysis of Child Mortality: Evidence from Nigeria. *Journal of Demography and Social Statistic*s, 1, 10-120.

Adeyemi R. A., Zewotir T. & Ramroop S. (2016). Semiparametric Multinomial Ordinal Model to Analyze Spatial Patterns of Child birth in Nigeria *International Journal of Environmental Research and public Health, 13, 1145*

Adler C.J. & Donlon D. (2020). Sexual dimorphism in deciduous crown traits of a European derived Australian sample. Forensic Sci Int 199: 29-37

Amzat, K. T. & Adeosun, S. A. (2014). On A Sequential Probit Model of Infant Mortality in Nigeria. *International Journal of Mathematics and Statistics Invention (IJMSI), 2(3), 89-94*

Balogun, O.S., Balogun, M.A., Abdulkadir, S.S. & Jibasen, D. (2014). A Comparison of the performance of Discriminant Analysis and the Logistic Regression methods in Classification of Drug Offenders in Kwara State. *International Journal of Advanced Research*, 2(10), 280-286.

Beki (2012). The Use of Discriminant Model and Logistic Regression for Tracking the Incidence of Broncho-Pulmonary Dysplasia among Infants. A dissertation submitted to Usman Danfodio University, Sokoto Nigeria.

Cornfield, A. (2010). Binary logistic model for predicting Myocardia infarction*: American journal of public health,* 81,15-21.

Danbaba A., Audu M. & Mahmud A. M. (2013). Investigation of risk factors of low birth weight using multivariate logistic regression analysis*: Journal of humanities, science and technology*. A publication of Niger State Polytechnic; 3(1), 65

Dattalo, P. (1995). A Comparison of Discriminant Analysis and Logistic regression: *Journal of Social Services Research*, 19(4), 121-144.

Edokpayi, A. A., Agho, C., Ezomo, J.E., Edosomwan, O.S. & Ogiugo, O.G. (2013). A Comparison of the Classification Performance of Discriminant Analysis and the Logistic Regression Methods in Identification of Oil Palm fruit Forms. A Paper Presented at the Annual Conference of Nigerian Statistical Association. 11-13th, September, 20-26.

Eneh, A.U. & Ugwu, R.O. (2021). the Proportion of Low Birth Weight Due to Small for Gestational age (SGA) and Prematurity. *The Internet Journal of Paediatrics and Neanatology*,13(1), 25

Erimafa J.T., Iduseri A.. & Edokpa I.W. (2009). Application of discriminant analysis to predict the class of degree for graduating students in a university system. *International Journal of Physical Sciences*. 4(1), 016-021.

Fernaudez, P., Abbas O., Dardenne, P., & Baeten, V. (2021). Discrimination of Corsican Honey by FT- Raman spectroscopy and Chenometrics. *Journal on Biotechnol Agron Soc Environ*, 15(1), 75- 84.

Herrera, J., Salmeron, B. & Hurtado, H. (2020). Prenatal Bio psychosocial risk Assessment and Low birth weight. *Social science medicine*. 44(8), 1107 – 1114.

Higgins, A. C, Mosley, J. E, Pencharz, P. B, Mikolainis, D. & Dubois, S. (1989). Impact of The Higgins Nutritional Intervention Program on Birth Weight a Within – Mother Analysis. *J Am Diet Assoc.* 89, 1097 – 103.

Hosmer D. W. & Lemeshow, S.. (1989): A review of goodness of fit statistics for use in the development of Logistic models*; American Journal Epidemiol*. 115, 92-106.

Hosmer, D.W. & Lemeshow, S. (2000). Applied Logistic Regression, (2nd ed.). Wiley and Son, inc. New York.

Jacdonmi, I., Suhainizam, M. S. & G. R. Jacdonmi, G. R. (2016). Breastfeeding, a child survival strategy against infant mortality in Nigeria. Current Science, 110(7), 1282-1293

Jiamwattanapong, K., Ingadapa, N, & Plubin, B. (2021). On Testing Homogeneity of Covariance Matrices with Box's M and the Approximate Tests for Multivariate Data. *European Journal of Applied Sciences*, 9(5), 426-436.

Kirkwood, R.B. & Stern. J.C.(2008). Essential Medical Statistics: (2nd ed.) Blackwell Science, London

Kleinbaum, .D.G, Kupper, .L.L, Muller, .K.E., & Morgens-Tern, H. (1982). Epidemiologic Research: Principles and Quantitative Methods. Van Nostrand Reinhold Company, New York, 281-417.

Kowsher Md, Nusrat jahan Prottasha, Anik Tahabilder, Kaiser Habib, Md Abdul-Rakib, Md Shammen Alam (2021). Predicting the appropriate Mode of Childbirth using Machine Learning Algorithm; *International Journal of advanced computer science and applications* 12(5), 86-97

Kramer, M.S. (2022). Balanced Protein/Energy Supplementation in Pregnancy. *Cochrane Database Sys Rev. 1(1), 15-24*

Liu, L., Kalter, H.D., Chu, Y., Kazmi, N., Koffi, A.K., Amouzou, A., Joos, O., Munos, M., & Black, R. E. (2016). Understanding Misclassification between Neonatal Deaths and Stillbirths: Empirical Evidence from Malawi. 11(12),1-11.

Mbanasor, J.A. & Nto, P.O.O., (2008), Discriminant Analysis of Livestock farmers' credit worthiness*: Journal of Nigeria Agriculture*. 1, 1-7

Mobil P, Landero A .De-Antoni G. Araujo Andraote C, Avila- Donoso H. & Moreno J. (2020). Multivariate Analysis of Raman Spectra applied to microbiology: Discriminant of Microorganism at the species level. 56 (5), 378-385.

Montgomery, N.E., White, M.E. & Martin, S.W. (1987). A Comparison of Discriminant analysis and Logistic Regression for the prediction of Coliform Mastitis in dairy Cows: *Canadian Journal of Veterinary Research*, 51(4), 495-498.

Ngwu, E. K. (2009). Birth weight of babies in relation to their mothers' nutritional knowledge and place of residence. *Nigerian Journal of Nutritional Science*. 30(2), 36 – 42.

Onyiriuka A. N. (2018), *Incidence of Low Birth Weight Infants in Twin Gestations.*Niger J Clin Pract; 13, 365 – 70.

Pathak, P., Kapil, U., Kapoor, S.K., Saxena, R., Kumar & A., Gupta, N. (2004) Prevalence of Multiple Micronutrient Deficiencies amongst Pregnant Women in rural area of Haryana. *Indian Pediatr*. 17(11), 1007 – 1014..

Rees, G., Brook, Z., Doyle, W. & Costeloe, K .(2005). The Nutritional Status of Women in theFirst Trimester of Pregnancy attending an inner-city Antenatal Department in the UK. *Journal of Royal Society for the Promotion of Health*, 125(5), 232 – 238.

Richardson, K. E., Innocent, A. I., & Okereke, O. S. (2014). Relationship between Household Income and Child Mortality in Nigeria. *American Journal of Life Sciences. Special Issue: Science, Society and Policy: Driving Towards Utopia or Dystopia, 2(6-4, 1-12.*

Usman, A. (2011). An Estimation Formula for Resonance Frequency using Sex and Height for healthy individuals and patients with incurable cancers. *Statistical methods for Biometric & Medical research*. Kaduna, Millennium Publisher (New Nigeria Newspaper).

Uthman, O. A. (2008). Effect of Low Birth Weight on Infant Mortality*; The Internet Journal of Epidemiology*, 6(1), 25-30

Vishwa N. M, Madaki U. Y, Vijay V. S & Babagana M. (2015). Application of Discriminant Analysis on Bronch-Plumonary Dysplasia among infants: *American Journal of Theoretical and Applied Statistics*, 4(2), 44-51

Wald, A. (1943) Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Transaction of the American Mathematical Society*, 54, 426-482 .

Wasowicz, W., Wolkanin, P., Bednarski, M., Gromadzinska, J., Sklodowska, M. & Grzybowska, K. (1993) Plasma trace element (Se, Zn, Cu) Concentrations on Maternal and umbilical cord blood in Poland: Relation with Birth weight, Gestation age, Parity, 38, 205 - 215

WHO (2020) Antenatal Care in Developing Countries: Promises, Achievements and Missed Opportunities – An Analysis of Trend Levels and Differentials WHO Geneva, *Bulletin of World Health Organisation*.

WHO & UNICEF (2004). Low Birth Weight, Country, Regional and Global Estimates

World Health Organisation Division of Family Health. The incidence of low birth weight. A critical review of available information. *World Health Statistics Quarterly*, (1980), 33, 197 – 224.

World Health Organization (2004). *International Statistical Classification of disease and related health problem*s, tenth revision (2nd ed)

World Health Organization (2020) Report of the Expert Committee on Maternal and Child Health. *Technical Report Series* No. 217

Yakubu, Y., Ahmed, S.S., Audu, I. & Usman, A. (2019). Binary logistic regression methods for modeling broncho-pneumonia status in infants from tertiary health institutions in north central Nigeria; *Journal of Applied Sciences and Environmental Management* 23(8), 1607-1614,

Yarnold, P.R., Hart, L.A., & Soltysik, R.C. (1994). Optimizing the Classification Performance of Logistic Regression and Fisher's Discriminant Analysis: *Journal of Educational and Psychological Measurement*, 54, 73-85

APPENDIX

S/NO	MH	MW	MA	BW	S	GA	P
1	1.65	68.2	30	1.5	M	34	2
2	1.67	67	30	2.2	F	32	4
3	1.74	65	32	2.7	F	34	5
4	1.51	50	25	2.1	M	30	3
5	1.54	59	25	2.7	M	36	0
6	1.55	56	33	1.5	M	36	0
7	1.65	65	27	2.5	M	34	2
8	1.67	65	20	2.2	M	34	4
9	1.52	52	30	2.5	M	34	0
10	1.62	60	26	2.8	F	36	0
11	1.76	51	22	2.7	M	36	1
12	1.6	68	28	1.5	F	34	5
13	1.16	60	25	3	F	36	1
14	1.71	46	26	2.7	F	32	6
15	1.59	57	38	3.5	F	36	0
16	1.45	73	25	3	M	36	1
17	1.65	64	37	2.7	M	36	6
18	1.75	52	28	3.2	M	36	5
19	1.55	55	26	2.7	M	36	2
20	1.7	75	34	3.1	M	36	2
21	1.54	58	23	2.1	F	32	1
22	1.66	64	21	2.8	F	36	0
23	1.7	52	29	2.8	M	38	7
24	1.58	58	23	4.1	M	36	4
25	1.54	55	25	1.3	M	36	0
26	1.69	56	34	3.7	M	36	0
27	1.5	54	22	2.5	F	34	2
28	1.55	56	28	2.5	M	36	2
29	1.65	63	25	3.2	M	32	1
30	1.6	70	35	2.2	M	32	4
31	1.58	50	30	1.8	M	34	1
32	1.5	60	29	2.3	M	36	7
33	1.56	68	30	2.5	F	34	0
34	1.52	45	32	3	F	32	5
35	1.57	55	25	3.5	F	36	4
36	1.52	52	22	3.2	M	32	4
37	1.67	62	30	3.7	M	34	5
38	1.6	60	30	3.6	F	34	0
39	1.45	59	29	2.3	F	32	1
40	1.64	65	22	3.2	M	32	4
41	1.56	50	18	2.9	F	32	2
42	1.58	56	34	2.5	M	30	3

43	1.6	60	28	3	M	26	0
44	1.51	50	30	2.1	M	32	5
45	1.74	74	22	3.5	F	36	0
46	1.78	63	35	3.2	M	36	5
47	1.76	75	35	2.3	M	36	0
48	1.5	50	36	3.9	M	32	5
49	1.62	61	34	2.5	M	30	2
50	1.6	73	32	4	F	36	1
51	1.44	45	31	2.2	F	40	0
52	1.5	45	28	2.2	M	40	2
53	1.8	60	31	3.2	M	36	4
54	1.55	58	26	2.8	M	36	5
55	1.7	50	28	1.7	M	38	0
56	1.64	68	20	2.8	M	38	0
57	1.52	45	23	2.5	F	30	0
58	1.54	50	35	3.5	F	28	3
59	1.7	60	28	3	F	26	0
60	1.62	62	34	2.2	M	30	3
61	1.6	61	27	2	M	30	2
62	1.54	53	27	2.8	F	36	2
63	1.56	50	31	2.5	F	32	3
64	1.65	65	22	3.5	M	32	4
65	1.84	80	25	3.5	M	32	4
66	1.6	61.5	18	1.5	M	32	3
67	1.5	50	22	4	F	32	2
68	1.6	69	30	2	M	32	0
69	1.65	65	23	4.5	F	34	0
70	1.5	52	25	2.3	M	36	0
71	1.79	69	23	2.5	M	36	5
72	1.6	60	33	4	F	30	6
73	1.65	60	28	3	F	32	5
74	1.55	57	30	2.3	M	36	2
75	1.74	60	24	3.2	F	32	3
76	1.6	60	30	3.2	F	36	4
77	1.45	60	22	2.2	F	36	4
78	1.6	58	22	2.2	M	32	2
79	1.62	55	30	2.5	M	36	1
80	1.52	45	25	2	M	36	1
81	1.45	58	35	2.4	M	36	1
82	1.6	55	35	2.5	M	36	7
83	1.69	56	35	2.9	M	32	7
84	1.45	59	30	2	M	34	8
85	1.64	60	30	2.5	M	36	1
86	1.45	60	29	2.2	M	32	1
87	1.52	52	30	3.2	F	34	0

88	1.65	64.5	30	3.2	F	34	0
89	1.69	64.1	25	2.8	F	36	1
90	1.66	65	33	3	F	36	0
91	1.65	64	30	3.2	F	31	4
92	1.55	54	31	2.5	M	31	6
93	1.6	60	27	2.2	F	30	3
94	1.6	45	21	2	F	34	0
95	1.8	70	29	3.3	F	32	3
96	1.6	60	35	3.5	M	36	0
97	1.45	60	28	2.2	F	34	2
98	1.48	60	30	2	M	34	4
99	1.52	52	25	2.1	F	34	6
100	1.65	45	28	1.8	M	32	2
101	1.6	60	26	2.2	F	36	1
102	1.63	68	21	4.3	F	30	2
103	1.6	50	21	3.2	M	32	5
104	1.47	72	33	2.7	M	36	5
105	1.44	40	30	3	F	36	3
106	1.55	54	20	3.4	F	34	2
107	1.45	54	29	2.8	M	36	4
108	1.53	51.9	25	1.9	F	32	2
109	1.7	70	20	3.9	M	34	5
110	1.65	65	27	1.7	F	34	4
111	1.5	50	29	2.9	M	32	4
112	1.62	62	36	2.2	M	32	4
113	1.57	65	21	3.4	M	32	3
114	1.59	61	24	2.3	F	30	4
115	1.7	62	29	2.5	F	26	5
116	1.75	60	37	3.6	F	32	7
117	1.6	62	32	2.4	M	32	6
118	1.52	56	18	3.4	M	32	5
119	1.79	45	22	2.2	M	30	1
120	1.52	53	24	4	F	36	0
121	1.75	76	29	3	M	36	1
122	1.64	65	30	2	F	32	1
123	1.5	60	40	2.5	F	36	1
124	1.3	50	34	2.5	M	36	1
125	1.7	61	36	2.4	M	34	8
126	1.6	60	22	3	F	34	7
127	1.45	48	22	2.1	F	36	7
128	1.48	62	26	2.8	M	34	1
129	1.52	52	19	3.2	F	36	3
130	1.5	45	25	2.1	M	38	2
131	1.7	70	34	2.5	F	40	2
132	1.46	46	35	2.5	F	26	4

133	1.7	78	28	2.2	M	32	4
134	1.65	62	33	3	F	32	3
135	1.69	65	30	2.5	M	34	2
136	1.83	50	19	2.2	M	36	5
137	1.6	48	30	2.5	F	34	6
138	1.7	70	25	3.2	M	40	0
139	1.6	60	27	2	F	40	2
140	1.69	65	35	2.2	F	36	0
141	1.65	65	22	3.5	F	34	2
142	1.65	68	30	2.1	M	36	4
143	1.51	52	25	2.8	F	36	4
144	1.55	57	30	3	F	36	3
145	1.41	45	24	3	M	32	2
146	1.5	62	19	3.2	F	34	2
147	1.53	71	30	3.5	F	34	3
148	1.62	64	27	4	M	40	4
149	1.59	54	30	3.6	M	32	7
150	1.6	60	30	2.5	F	32	1
151	1.61	60	25	3.1	F	36	4
152	1.44	45	36	2	F	30	1
153	1.6	62	20	3.8	F	30	2
154	1.64	60	35	3	F	34	2
155	1.53	34	25	3	M	38	0
156	1.5	60	35	2.5	M	32	0
157	1.61	65	27	3.5	M	34	4
158	1.52	52	26	2.5	F	34	7
159	1.54	60	20	2.1	M	36	0
160	1.58	72	40	2.8	F	36	1
161	1.6	68	20	1.5	M	32	2
162	1.54	50	28	2.5	M	36	2
163	1.65	65	30	3.1	F	38	5
164	1.57	42	25	2.3	M	34	6
165	1.6	65	20	3.3	F	34	1
166	1.65	80	23	3.8	M	32	0
167	1.63	62	20	2.3	F	32	6
168	1.7	72	25	3.5	M	28	1
169	1.71	71	25	3.7	M	36	5
170	1.6	57	27	2.3	F	36	0
171	1.5	70	25	3	F	36	0
172	1.6	63.5	19	1.8	M	36	4
173	1.6	60	25	2.3	F	36	2
174	1.5	50	27	3.1	F	28	0
175	1.45	60	27	2.2	M	36	3
176	1.63	54	24	2.5	F	36	5
177	1.47	43	29	2.5	M	36	2

178	1.45	51.6	27	3.5	F	36	4
179	1.45	47	32	2.4	F	36	2
180	1.6	60	30	2.5	M	36	3
181	1.79	77	29	3.8	M	36	0
182	1.6	70	27	3	F	28	0
183	1.61	60	30	2	M	36	2
184	1.5	53	30	2	M	36	4
185	1.8	80	22	2.5	F	36	0
186	1.59	42	25	2	F	36	0
187	1.7	50	17	1.7	F	36	1
188	1.55	65	17	3.7	F	36	6
189	1.64	65	23	1.2	M	38	6
190	1.55	55	38	2	M	36	1
191	1.62	60	28	3	M	36	6
192	1.5	56	30	3.5	M	36	5
193	1.73	60	20	2	F	34	2
194	1.64	74	35	2.3	F	38	2
195	1.8	80	20	3	F	36	1
196	1.6	65	20	2.9	F	36	0
197	1.61	61	28	2.7	M	36	2
198	1.55	55	21	3	M	36	2
199	1.46	52	28	1.5	F	32	4
200	1.53	55	34	3.2	M	36	1
201	1.61	61	30	2.8	F	34	2
202	1.64	65	30	2	F	32	4
203	1.67	65	32	4	M	34	5
204	1.6	65	25	2	M	30	3
205	1.59	57	25	3.5	M	36	0
206	1.54	57	33	3.2	M	36	0
207	1.53	53	27	2.2	M	34	2
208	1.61	60	20	2.2	M	34	4
209	1.5	52	30	2.6	F	34	0
210	1.58	58	26	2	M	36	0
211	1.46	45	22	2	F	36	1
212	1.74	78	28	3.8	F	34	5
213	1.52	58	25	1.8	F	36	1
214	1.57	60	26	3.7	F	32	6
215	1.64	60	38	2.4	M	36	0
216	1.6	77	25	4	M	36	1
217	1.73	76	37	3	M	36	6
218	1.7	50	28	2.2	M	36	5
219	1.5	60	26	3	M	36	2
220	1.62	60	34	3.2	F	36	2
221	1.71	60	23	2	F	32	1
222	1.84	74	21	2.5	M	36	0

223	1.55	53	29	3.2	M	38	7
224	1.52	52	23	3.3	M	36	4
225	1.54	60	25	2.5	M	36	0
226	1.56	65	34	3.4	F	36	0
227	1.55	55	22	2.3	M	34	2
228	1.72	73	28	4	M	36	2
229	1.67	65	25	1.7	M	32	1
230	1.46	45	35	3.4	M	32	4
231	1.5	52	30	2.8	M	34	1
232	1.52	52	29	2.8	F	36	7
233	1.48	72	30	4	F	34	0
234	1.72	65	32	1.7	F	32	5
235	1.4	55	25	2.2	M	36	4
236	1.6	53	22	3.5	M	32	4
237	1.55	57	30	3	F	34	5
238	1.48	53	30	3.5	F	34	0
239	1.45	60	29	3	M	32	1
240	1.4	48	22	3.5	F	32	4
241	1.8	70	18	2.1	M	32	2
242	1.45	45	34	2.6	M	30	3
243	1.6	50	28	2.8	M	26	0
244	1.69	65	30	3.1	F	32	5
245	1.59	59	22	3.3	M	36	0
246	1.72	61	35	3.1	M	36	5
247	1.62	65	35	2.4	M	36	0
248	1.61	63	36	3	M	32	5
249	1.71	70	34	4	F	30	2
250	1.56	72	32	4	F	36	1
251	1.69	65	31	2	M	40	0
252	1.6	60	28	3.8	M	40	2
253	1.55	60	31	3.4	M	36	4
254	1.68	68	26	2.5	M	36	5
255	1.6	55	28	2.3	M	38	0
256	1.6	63	20	3.4	F	38	0
257	1.64	62	23	3.5	F	30	0
258	1.5	68	35	2.5	F	28	3
259	1.78	55	28	2.3	M	26	0
260	1.52	49	34	2.3	M	30	3
261	1.46	42	27	3.5	M	30	2
262	1.72	72	27	2.1	F	36	2
263	1.48	60	31	2.1	M	32	3
264	1.7	75	22	3.1	M	32	4
265	1.53	50	25	2.8	M	32	4
266	1.47	60	18	3.1	F	32	3
267	1.55	55	22	2.2	M	32	2

268	1.52	50	30	3.2	F	32	0
269	1.7	72.3	23	3	M	34	0
270	1.6	80	25	3	M	36	0
271	1.45	49	23	3	F	36	5
272	1.65	63	33	3.8	F	30	6
273	1.69	59	28	3.2	M	32	5
274	1.6	60	30	2.1	F	36	2
275	1.5	65.1	24	2.5	F	32	3
276	1.77	60	30	2.7	F	36	4
277	1.6	61	22	2	M	36	4
278	1.49	60	22	2.6	M	32	2
279	1.7	70	30	3.5	M	36	1
280	1.62	60	25	2.5	M	36	1
281	1.45	60	35	2.2	M	36	1
282	1.5	50	35	3.7	M	36	7
283	1.45	45	35	2.1	M	32	7
284	1.75	80	30	3.1	M	34	8
285	1.7	67	30	2.5	M	36	1
286	1.72	70	29	2.3	F	32	1
287	1.64	60	30	3.2	F	34	0
288	1.52	48	30	2	F	34	0
289	1.7	70	25	3.5	F	36	1
290	1.56	58	33	2.1	F	36	0
291	1.7	70	30	1.2	M	31	4
292	1.5	60	31	3.5	F	31	6
293	1.47	60	27	2.2	F	30	3
294	1.59	50	21	2.3	F	34	0
295	1.64	65	29	2	M	32	3
296	1.67	67	35	2.3	F	36	0
297	1.61	62	28	3.8	M	34	2
298	1.5	50	30	2.8	F	34	4
299	1.5	60	25	1.7	M	34	6
300	1.7	68	28	3.1	F	32	2
301	1.86	50	26	1.8	F	36	1
302	1.75	50	21	2.3	M	30	2
303	1.79	45	21	2.1	M	32	5
304	1.55	53	33	3.2	F	36	5
305	1.6	60	30	3.9	F	36	3
306	1.74	60	20	3.2	M	34	2
307	1.45	45	29	2	F	36	4
308	1.61	46	25	2.5	M	32	2
309	1.45	45	20	3.5	F	34	5
310	1.48	46	27	2.4	M	34	4
311	1.64	65	29	2.5	M	32	4
312	1.66	70	36	3.5	M	32	4

313	1.7	67	21	2.9	F	32	3
314	1.6	60	24	2.2	F	30	4
315	1.64	52	29	3.5	F	26	5
316	1.5	50	37	3.2	M	32	7
317	1.6	50	32	3.2	M	32	6
318	1.75	58	18	3	M	32	5
319	1.7	70	22	3	F	30	1
320	1.64	60	24	4.5	M	36	0
321	1.7	70	29	3.6	F	36	1
322	1.42	40	30	3.5	F	32	1
323	1.6	50	40	2.8	M	36	1
324	1.48	50	34	2.5	M	36	1
325	1.52	49	36	2.3	F	34	8
326	1.78	47	22	2.3	F	34	7
327	1.45	55	22	2.7	M	36	7
328	1.49	45	26	2.1	F	34	1
329	1.54	50	19	2.1	M	36	3
330	1.62	60	25	2.1	F	38	2
331	1.73	70	34	2.9	F	40	2
332	1.7	70	35	3.5	M	26	4
333	1.62	60	28	2.5	F	32	4
334	1.5	50	33	3.5	M	32	3
335	1.54	70	30	3.5	M	34	2
336	1.53	60	19	3.2	F	36	5
337	1.65	66	30	3	M	34	6
338	1.52	70	25	3.3	F	40	0
339	1.45	60	27	2.2	F	40	2
340	1.6	60	35	2.8	F	36	0
341	1.46	44	22	1.9	M	34	2
342	1.44	42	30	2.1	F	36	4
343	1.48	49	25	2.9	F	36	4
344	1.61	60	30	2.5	M	36	3
345	1.59	50	24	3.8	F	32	2
346	1.6	60	19	3	F	34	2
347	1.65	65	30	3.7	M	34	3
348	1.78	42	27	2.8	M	40	4
349	1.74	60	30	2.5	F	32	7
350	1.56	59	30	2.7	F	32	1
351	1.6	60	25	2	F	36	4
352	1.79	57	36	3.1	F	30	1
353	1.6	60	20	2.1	F	30	2
354	1.64	65	35	3	M	34	2
355	1.49	49	25	2	M	38	0
356	1.71	65	35	2.2	M	32	0
357	1.72	56	27	3.2	F	34	4

358	1.64	60	26	3.4	M	34	7
359	1.52	50	20	2.7	F	36	0
360	1.75	78	40	2.9	M	36	1
361	1.71	79	20	2.9	M	32	2
362	1.59	59	28	2.4	F	36	2
363	1.59	54	30	2.5	M	38	5
364	1.53	54	25	2.8	F	34	6
365	1.52	50	20	2.8	M	34	1
366	1.53	58	23	2.5	F	32	0
367	1.64	64.2	20	3.2	M	32	6
368	1.6	63	25	3.2	M	28	1
369	1.7	63	25	2.5	F	36	5
370	1.45	46	27	3.2	F	36	0
371	1.52	51	25	2.3	M	36	0
372	1.59	57	19	1.9	F	36	4
373	1.46	45	25	3	F	36	2
374	1.61	64	27	2.2	M	28	0
375	1.65	60	27	2.1	F	36	3
376	1.71	51	24	2.7	M	36	5
377	1.52	53	29	2.1	F	36	2
378	1.59	60	27	3	F	36	4
379	1.78	100	32	2.2	M	36	2
380	1.44	50	30	3	M	36	3
381	1.6	58	29	2.5	F	36	0
382	1.72	70	27	3	M	28	0
383	1.45	45	30	2.6	M	36	2
384	1.6	45	30	2.4	F	36	4
385	1.62	53	22	2	F	36	0
386	1.75	80	25	3.2	F	36	0
387	1.59	59	17	3	F	36	1
388	1.57	55	17	1.6	M	36	6
389	1.46	46	23	2.2	M	38	6
390	1.78	79	38	4	M	36	1
391	1.45	60	28	2.2	M	36	6
392	1.6	59	30	2.8	F	36	5
393	1.57	57	20	1.6	F	34	2
394	1.69	60	35	2.1	F	38	2
395	1.6	60	20	2.1	F	36	1
396	1.65	53	20	2.4	M	36	0
397	1.6	60	28	2.5	M	36	2
398	1.75	70	21	2.5	F	36	2
399	1.5	56	28	1.5	M	32	4
400	1.54	65	34	3.4	F	36	1
401	1.45	45	30	2.5	F	34	2
402	1.48	54	30	2.8	M	32	4

403	1.6	58	32	3.5	M	34	5
404	1.55	65	25	3.7	M	30	3
405	1.68	65	25	2.1	M	36	0
406	1.6	68	33	1.5	M	36	0
407	1.62	60	27	2.2	M	34	2
408	1.6	59	20	2.2	F	34	4
409	1.65	75	30	2	M	34	0
410	1.5	76	26	3.5	F	36	0
411	1.78	59	22	3.2	F	36	1
412	1.64	52	28	2.3	F	34	5
413	1.45	46	25	2.5	F	36	1
414	1.5	50	26	3.2	M	32	6
415	1.71	79	38	2.9	M	36	0
416	1.78	73	25	3.2	M	36	1
417	1.6	64	37	2.5	M	36	6
418	1.65	65	28	1.7	M	36	5
419	1.43	40	26	2.6	F	36	2
420	1.7	72	34	3.8	F	36	2
421	1.62	57	23	2.2	M	32	1
422	1.78	64	21	2	M	36	0
423	1.6	58	29	3.5	M	38	7
424	1.5	56	23	3.5	M	36	4
425	1.65	61	25	3.6	F	36	0
426	1.47	50	34	3	M	36	0
427	1.49	47	22	2.3	M	34	2
428	1.44	50	28	2	M	36	2
429	1.6	65	25	2.1	M	32	1
430	1.65	66	35	2.8	M	32	4
431	1.7	60	30	2.5	F	34	1
432	1.54	68	29	3.5	F	36	7
433	1.6	60	30	2.2	F	34	0
434	1.6	55	32	2.6	M	32	5
435	1.71	72	25	3	M	36	4
436	1.68	65	22	2.1	F	32	4
437	1.7	50	30	2.3	F	34	5
438	1.62	50	30	1.8	M	34	0
439	1.58	65	29	2.5	F	32	1
440	1.5	49	22	2.1	M	32	4
441	1.51	50	18	2.2	M	32	2
442	1.72	70	34	3	M	30	3
443	1.62	58	28	2.9	F	26	0
444	1.74	70	30	2.8	M	32	5
445	1.45	45	22	2.9	M	36	0
446	1.6	60	35	3.2	M	36	5
447	1.8	82	35	3.5	M	36	0

448	1.6	63	36	2.1	F	32	5
449	1.6	64	34	3.5	F	30	2
450	1.62	65	32	2.8	M	36	1
451	1.61	60	31	4.1	M	40	0
452	1.84	84	28	2.5	M	40	2
453	1.6	60	31	2.5	M	36	4
454	1.5	50	26	3.2	M	36	5
455	1.49	40	28	2.5	F	38	0
456	1.55	65	20	2.4	F	38	0
457	1.53	60	23	2.1	F	30	0
458	1.79	85	35	3.2	M	28	3
459	1.77	64	28	2.3	M	26	0
460	1.6	68	34	3.5	M	30	3
461	1.65	68	27	2.8	F	30	2
462	1.65	70	30	3.1	M	36	2
463	1.45	60	30	2.1	M	32	3
464	1.51	55	32	1.9	M	32	4
465	1.46	45	25	3.4	F	32	4
466	1.6	60	25	3.5	M	32	3
467	1.45	60	33	2.2	F	32	2
468	1.44	60	27	2.8	M	32	0
469	1.51	56	20	4	M	34	0
470	1.59	55	30	2.3	F	36	0
471	1.48	53	26	3.5	F	36	5
472	1.52	41	22	2	M	30	6
473	1.57	58	28	2.8	F	32	5
474	1.6	58	25	3.6	F	36	2
475	1.59	87	26	3.5	F	32	3
476	1.6	44	38	2.5	M	36	4
477	1.47	46	25	1.5	M	36	4
478	1.45	40	37	2.2	M	32	2
479	1.6	57	28	2.2	M	36	1
480	1.7	50	26	3	M	36	1
481	1.55	56	34	2.9	M	36	1
482	1.52	59	23	3.4	M	36	7
483	1.62	60	21	2.2	M	32	7
484	1.55	55	29	3.4	M	34	8
485	1.8	70	23	3	F	36	1
486	1.59	51	25	2.4	F	32	1
487	1.49	50	34	3.2	F	34	0
488	1.64	60	22	2.3	F	34	0
489	1.64	63	28	2.4	F	36	1
490	1.53	70	25	3	M	36	0
491	1.58	62	35	2.8	F	31	4
492	1.6	60	30	2.2	F	31	6

493	1.52	50	29	2.4	F	30	3
494	1.54	50	30	3.2	M	34	0
495	1.6	60	32	3.2	F	32	3
496	1.71	70	25	3.3	M	36	0
497	1.63	62	22	3.5	F	34	2
498	1.62	65	30	2.8	M	34	4
499	1.43	56	30	3.8	F	34	6
500	1.69	60	29	2.3	F	32	2
501	1.65	61	22	3	M	36	1
502	1.65	67	18	3.2	M	30	2
503	1.6	59	34	2.1	F	32	5
504	1.54	61	28	2.8	F	36	5
505	1.69	63	30	2.9	M	36	3
506	1.44	48	22	2.2	F	34	2
507	1.57	58	35	3.4	M	36	4
508	1.6	60	35	3.6	F	32	2
509	1.68	67	36	3.1	M	34	5
510	1.55	58	34	3	M	34	4
511	1.53	54	32	4	M	32	4
512	1.67	40	31	3	F	32	4
513	1.65	65	28	3	F	32	3
514	1.66	56	31	2.3	F	30	4
515	1.5	68	26	3.1	M	26	5
516	1.63	63	28	3.7	M	32	7
517	1.58	45	20	4.5	M	32	6
518	1.71	54	23	2.5	F	32	5
519	1.65	50	35	2.9	M	30	1
520	1.69	68	28	4.8	F	36	0
521	1.56	58	34	2.8	F	36	1
522	1.6	60	27	3.2	M	32	1
523	1.59	59	27	1.9	M	36	1
524	1.45	60	31	2.1	F	36	1
525	1.5	56	22	2.3	F	34	8
526	1.55	58.3	25	2.5	M	34	7
527	1.6	60	18	3.5	F	36	7
528	1.63	65	22	2.2	M	34	1
529	1.67	50	30	2.5	F	36	3
530	1.59	60	23	2	F	38	2
531	1.71	70	25	4	M	40	2
532	1.72	59	23	2	F	26	4
533	1.6	60	33	3	M	32	4
534	1.5	65	28	3.1	M	32	3
535	1.58	69	30	2.1	F	34	2
536	1.55	45	24	2.5	M	36	5
537	1.5	50	30	4	F	34	6

538	1.6	60	22	2.1	F	40	0
539	1.61	61	22	2.8	F	40	2
540	1.5	50	30	2.3	M	36	0
541	1.52	40	25	2.3	F	34	2
542	1.68	56	35	2.5	F	36	4
543	1.63	64	35	2.7	M	36	4
544	1.7	70	35	3.5	F	36	3
545	1.81	80	30	3.5	F	32	2
546	1.52	60.5	30	3	M	34	2
547	1.65	65	29	2.2	M	34	3
548	1.66	65	30	3.5	F	40	4
549	1.44	60	30	2.1	F	32	7
550	1.49	46	25	2.5	F	32	1
551	1.44	43	33	2.9	F	36	4
552	1.47	50	30	2.5	F	30	1
553	1.45	60	31	2.1	M	30	2
554	1.6	53	27	1.8	M	34	2
555	1.57	55	21	2.5	M	38	0
556	1.5	50	29	2.5	F	32	0
557	1.53	53	35	4.2	M	34	4
558	1.5	63	28	2.3	F	34	7
559	1.61	60	30	2.5	M	36	0
560	1.53	50	25	3.5	M	36	1
561	1.51	52	28	2.5	F	32	2
562	1.82	46	26	2.5	M	36	2
563	1.71	60	21	2.2	F	38	5
564	1.42	62	21	2.1	M	34	6
565	1.5	63	33	2.3	F	34	1
566	1.6	60	30	2.5	M	32	0
567	1.58	65	20	2.7	M	32	6
568	1.71	63	29	2.4	F	28	1
569	1.59	52	25	2.8	F	36	5
570	1.44	50	20	2	M	36	0
571	1.45	60	27	2.1	F	36	0
572	1.6	60	29	4	F	36	4
573	1.9	87	36	3.7	M	36	2
574	1.59	68	21	3	F	28	0
575	1.65	66.3	24	2.4	M	36	3
576	1.68	68	29	2.5	F	36	5
577	1.58	85	37	2.5	F	36	2
578	1.57	50	32	2.4	M	36	4
579	1.64	63	18	2.5	M	36	2
580	1.63	60	22	3.5	F	36	3
581	1.56	50	24	2.5	M	36	0
582	1.59	54	29	3.5	M	28	0

583	1.5	50	30	3.2	F	36	2
584	1.62	54	40	2.3	F	36	4
585	1.69	50	34	2.5	F	36	0
586	1.45	61	36	2.1	F	36	0
587	1.53	53	22	3.5	M	36	1
588	1.52	58	22	1.8	M	36	6
589	1.7	80	26	3.7	M	38	6
590	1.48	87	19	2.7	M	36	1
591	1.59	55	25	2.5	F	36	6
592	1.7	66.4	34	3.5	F	36	5
593	1.64	60	35	2.9	F	34	2
594	1.69	70	28	2.5	F	38	2
595	1.6	62	33	2.1	M	36	1
596	1.46	40	30	2.5	M	36	0
597	1.71	50	19	2.1	F	36	2
598	1.6	61	30	2	F	36	2
599	1.62	65	25	2.8	F	32	4
600	1.67	60	27	2.3	F	36	1
601	1.6	63	35	2.2	F	32	4
602	1.46	63	22	3	M	32	6
603	1.45	44	30	3.8	F	34	6
604	1.61	62	25	3.5	M	30	5
605	1.64	60	30	3.6	M	32	3
606	1.45	85	24	3.5	F	30	2
607	1.75	75	19	3.5	M	36	4
608	1.43	60	30	2.5	F	34	1

Source: Jummai Babangida Maternal and Neonatal Hospital Minna
Where **MH** = Maternal height, **MW** = Maternal Weight, **MA** = Maternal Age, BW = Birth Weight, **S** = Sex, **GA** = Gestational Age and **P** = Parity