MODELLING THE LINKAGE BETWEEN DIETARY PATTERNS AND UNDERNUTRITION OUTCOMES IN UNDER FIVE CHILDREN IN MALAWI

MASTER OF SCIENCE (BIOSTATISTICS) THESIS

TAPIWA YVONNE MACHINJIRI

UNIVERSITY OF MALAWI

MAY, 2024

MODELLING THE LINKAGE BETWEEN DIETARY PATTERNS AND UNDERNUTRITION OUTCOMES IN UNDER FIVE CHILDREN IN MALAWI

MSc (BIOSTATISTICS) THESIS

By

TAPIWA YVONNE MACHINJIRI

Bachelor of Science-University of Malawi (The Malawi Polytechnic)

Thesis submitted to the Department of Mathematical Sciences, Faculty of Science, in partial fulfilment of the requirement for degree of Master of Science (Biostatistics)

UNIVERSITY OF MALAWI

MAY, 2024

DECLARATION

I, the undersigned hereby declare that this thesis is my own original work which has not been submitted to any other institution for similar purposes. Where other people's work has been used acknowledgements have been made.

By

Tapiwa Yvonne Machinjiri

Full Legal Name

Signature

Date

MAY, 2024

CERTIFICATE OF APPROVAL

The undersigned certify that this thesis represents the student's own work and effort and has been submitted with our approval.

Signature:―――――――――――――――― Date:――――――――――――

LAWRENCE KAZEMBE, PhD (Professor)

Supervisor

Signature:―――――――――――――――― Date:――――――――――――

MICHAEL CHIPETA, PhD

Co-Supervisor

DEDICATION

To God almighty for making it possible in every way for me to achieve my goal.

Praise be to God.

ACKNOWLEDGEMENT

I sincerely thank my supervisors for their support. Firstly, to Prof. Lawrence Kazembe for the ideas and knowledge acquired so far, for without him, I couldn't make it this far. He was very supportive and so patient from the start. Am also grateful to Dr. Michael Chipeta for the support and encouragement, it gave me strength and courage in my work that I can make it. Many thanks to NSO, demography section for their support, providing me the data and helping me in understanding the same.

Special thanks also goes to all my lecturers who imparted biostatistics knowledge in me throughout my learning at Chancellor College. Mr. Pieter Nthenda, the Diocesan CADECOM Coordinator for Mangochi CADECOM who approved me to pursue further studies and all CADECOM Mangochi colleagues for their support and patience the whole time I was pursuing my studies and had to do some work on my behalf just to get things moving.

My sincere gratitude to my mother Faith Msuku, my aunt Doreen Msuku, my husband Happy Gama, my daughter Tsanzo Gama and my whole family for their moral, spiritual and material support during my studies. Your prayers and encouragement kept me working towards achieving my goal. My kind gratitude to Mrs Glory Mshali and family for their support during my stay in Zomba, thanks so much. God bless you all.

ABSTRACT

Undernutrition remains a major public health concern resulting from nutritional deficiences. In Malawi 37% of underfive children are stunted, 3% are wasted and 12% are underweight. This study focused on modelling the linkage between dietary patterns and undernutrition outcomes in underfive children in Malawi. The study used data from 2015-16 Malawi Demographic and Health Survey which included a 24 hour recall of food consumed by underfive children and anthropometric measurements. Principal component regression (PCR) and reduced rank regression (RRR) were used to derive dietary patterns and binary logistic regression was used to model the linkage. Results showed that carbohydrates-legumes-fruits-vegetables-animal protein-fats-sugar dietary pattern derived from PCR was associated with stunting at $p<0.05$ (medium DDL OR=1.483, $p < 0.05$, cI: 1.218-1.806 and high DDL (OR=1.298, $p < 0.05$, CI: 1.063-1.586). Furthermore, dietary patterns derived from RRR namely carbohydrates-fruits-vegetables-fish pattern , Legumes-animal protein-sugar pattern and solid-foods-other-liquid pattern were associated with stunting, wasting and underweight with atleast one DDL having $p < 0.05$. Also, demographic factors including education level, region, residence and wealth index were associated with undernutrition at $p < 0.05$. The current study support consideration of dietary patterns and demographic factors in combating undernutrition. Therefore, stakeholders should implement nutrition-sensitive interventions promoting balanced diet with consideration of demographic factors for the health well being of underfive children and Malawi's economic progress.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

Abbreviation		Description
AIIC	:	Akaike Information Criterion
DDL	:	Dietary Diversity Level
IYCF	:	Infant and Young Child Feeding
MDGs	:	Malawi Development Goals
MDHS	:	Malawi Demographic and Health Survey
MNS	:	Micronutrient Survey
NCD	:	Non Communicable Diseases
NMNP	:	National Multi-sectoral Nutritional Policy
NSO	:	National Statistical Office
PCA	:	Principal Component Analysis
PCR	:	Principal Component Regression
RRR	:	Reduced Rank Regression
SDGs	:	Sustainable Development Goals
SUN	:	Scalling Up Nutrition
SD	:	Standard Deviation
UN	:	United Nations
WHO	:	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 Background

Undernutrition refers to the ”insufficient intake of energy and nutrients to meet an individual’s needs to maintain good health” (Maleta, 2006). According to Yue, Zhang, Li, and Qin (2022) undernutrition remains a major public health concern resulting from nutritional deficiences and its prevalence remain high world wide .

1.2 Burden and Distribution of Undernutrition Outcomes

The Global Nutrition Report (2020) states that the global burden of undernutrition is high whilst on the other hand progress made to combat it is low. According to the Global Nutrition Report (2021), undernutrition causes multiple burdens to under five children, 149 million children are stunted, 45 million are wasted and 30 million babies are underweight . In Africa, 30.7% of underfive children are stunted thus higher than the global average of 22% , 6% are wasted and 13.7% are underweight (G. N. Report, 2020). Undernutrition especially in low social economic countries remains a big problem (Mphamba, Chirwa, & Mazalale, 2024). Yue et al. (2022) explains that undernutrition contributes to childhood death, around 45 % of deaths for underfive children are linked to nutrition related factors. Undernutrition also delays development of global countries (Nunget, Levin, Harry, & Hutchinson, 2019). In general, the cost of addressing poor diets and undernutrition is high, the world needs 10.8 billion United States dollars as additional funding between 2022 and 2030 to meet global nutrition targets which includes childhood stunting, wasting and underweight (G. N. Report, 2020).

In Africa, distribution of nutritional outcomes mostly depends on background characteristics such as mother's education, geograhical location, economic status and biological factors (Tesfay, Javanparast, Gesesew, Mwanri, & Ziersch, 2022). Specifically, in Malawi, stunting is higher among children residing in rural areas (39%) compared to those in urban areas (25%). Stunting levels are higher for children born from mothers with no education (42%)as compared to those born from mothers with secondary education (12%) (Mphamba et al., 2024). Prevalence of stuntedness in children born from wealthy families is less (24%) than those from poor families (46%) (Profile, 2018). Furthermore prevalence of nutritional outcomes also depend on biological factors such as child's age, child's weight at birth and mother's body mass index (BMI) (G. N. Report, 2020). Fifty percent (50%) of children reported to have low birth weight are stunted. Children born from mothers with BMI less than 18.5 are more likely to be stunted, wasted, or underweight compared to children born to mothers with normal BMI or those who are overweight/obese (Machira & Chirwa, 2020).

The burdens of poor nutritional outcomes cannot be over emphasized in Malawi. Malnutrition in childhood results in a number of adverse short term and long term consequences for well-being and survival of children (Tesfay et al., 2022). Malawi being one of the developing countries in which 66% of the people spend less than $1.9 a day undernutrition in children remains a public health problem . Underweight, stunting and wasting have significant consequences on economic productivity, human capital and national development.

There are a number of social economic impacts of child under-nutrition, these include health care costs, additional burdens to education system, low productivity by workforce (Nunget et al., 2019). About 16 billion kwacha is lost annually due to reduced productivity from those who suffered from stuntedness in their childhood. In 2012, 10.3 % of GDP was lost due to under-nutrition (Guide, 2012). Furthermore, the burdens of poor nutrition range from health, education and human capital (Tesfay et al., 2022). The burdens include increased risk of specific health problems such as anaemia, diarrhoea and respiratory infections; high risk of dying; cognitive delays leading to high risk of repeating grades at school, 18% of grade repitions are experiences by stunted children; affects human capital and productivity for both manual and non-manual productivity, 60% of working age population were stunted when young. Undernutrition in children costs 35% of public budget allocated to the health sector in Malawi (Social, 2020) and (Nunget et al., 2019).

1.3 Undernutrition Outcomes in Malawi

Undernutrition is the most prominent nutritional outcome in under five children, it is one of the long existing outcomes of nutritional deficiences (Maleta, 2006). Poor diet is a major cause of undernutrition, increased suffering from infectious diseases such as Malaria also results in under nutrition. Undernutritional outomes in Malawi include underweight, stunting and wasting (Ntenda & Chuang, 2018). According to Kelly (2011) underweight (weight-for-age) is defined as weight of a child relative to the weight of a child of the same age in a reference population, children who have low weight-for-age are described as being 'underweight' (Mphamba et al., 2024). Secondly, stunting (height-for-age) is defined as height of a child relative to the height of a child of the same age in a reference population, children who have low height-for-age are described as being 'stunted'. Lastly,wasting (weight-for-height)is defined as weight of a child relative to the weight of a child of the same height in a reference population, children who have low weight-for-height are described as being 'wasted' (Machira & Chirwa, 2020).

The Malawi Nutrition Profile Report (2018) states that in Malawi 12% of the children are underweight, 37% are stunted and 3% are wasted. Furthermore, Tesfay et al. reports that undernutrition hasten progression of HIV and results in weakening of the immune system there by affecting the body's ability to fight and recover from illness .

1.4 Dietary Patterns and Undernutrition

Dietary patterns are defined as "quantities, proportions, varieties or combination of different foods, drinks and nutrients in diets and the frequency with which they are habitually consumed" (Krebs-Smith & Hoffman, 2014). Eating a diverse diet that include foods from different food groups contributes to a healthy well being of a person (Grewall & Wakim, 2021). Globally, peoples' dietary groups consumption including those of underfive children depend on their social-economic status and location, thus whether in rural or urban areas. The Global Nutrition Report (2018) attributes undernutrition to poor feeding practices . The report states that in Africa only 65% of people in both rural and urban areas adhere to the minimum accepted diet and 69% of people follow the recommended minimum dietary diversity. The (2021) global nutrition report further explains that among children under 5 years of age 149 million are stunted and 45 million are wasted across the world. The rates of undernutrition are substantially higher in the Sub-Saharan African region compared to western countries .

A study done by Khamis, Mwanri, Ntwenya, and Kreppel (2019) in Tanzania to examine the influence of dietary diversity on the nutritional status of children aged 6 months to 23 months showed that children who consumed diverse diets were less likely to be undernourished than those who had a less diverse diet (Khamis et al., 2019). This shows that there is a link between feeding habits and undernutrition. World Health Organisation (WHO) recommends a minimum acceptable diet for children from 6 to 23 months of age. Minimum acceptable diet refers to feeding a child that is breastfed or not from at least four food groups from the recommended seven food groups which include grains, roots, and tubers; legumes and nuts; dairy products (milk, yoghurt, cheese); flesh foods (meat, fish, poultry, and liver or organ meat); eggs; vitamin A rich fruits and vegetables; and other fruits and vegetables (UNICEF & WHO, 2008).

In Malawi, dietary groups of underfive children depend on the social-economic status, child's age and breast feeding status. Children below the age of 6 months are recommended for breast-feeding since breast milk is enough to provide all necessary nutrients (Grewall & Wakim, 2021). On the other hand, from 6 months children are introduced to complementary feeding, they are fed a variety of foods to ensure that they consume all required nutrients (G. N. Report, 2021).
Machira and Chirwa (2020) conducted a study to examine dietary consumption and its effect on nutrition outcomes among under-five children in rural Malawi. The study found that dietary consumption has a link with undernutrion outcomes thus underweight, stunting and wasting. They recommdeded the need for education advocacy targeting parents on exclusive breastfeeding and consistency in dietary foods given to children inorder to combat undernutrition.

The Department of Nutrition, HIV and AIDS (2018) report show that 25% of breast-fed children consumed appropriate food groups. However, only 8% of children from 6 to 23 months of age meet minimum standards for all Infant and Young Children Feeding (IYCF) practices. The percentage of children meeting minimum diet increase with mothers education (no education 4%, more than secondary education 24%), background (breastfed 24%, non-breast fed 30%) and geographical areas (urban 43%, rural 22%) (Nyanhanda, Mwanri, & Mude, 2023).

Statistical analyses of linking diets or foods to undernutrition have been done worldwide, however, it should be recognized that people do not only eat certain foods or nutrients but a complex mixture of foods hence the need to use dietary approach. This approach was initially proposed at the White House Conference on Food, Nutrition and Health in 1969 (Schwerin et al., 1981). Since then a number of researchers have been using dietary approach in linking diets to outcomes such as undernutrition and illnesses.

Studies are required to assess linkages that exist between dietary patterns and undernutrition. Principal component regression and reduced rank regression are dimension reduction statistical methods widely used in modelling the linkage beween dietary patterns and nutrition outcomes. Few researchers have applied PCR and RRR in Malawi to assess dietary patterns and their outcomes. This study used PCR and RRR to assess the linkage that exist between dietary patterns and nutritional outcomes in under-five children in Malawi.

1.5 Epidemiology and Control of Nutritional Outcomes

The high level panel of experts on nutrition and food systems (2017) report explains that globally one in every three persons is malnourished. The projected figure could reach to one in every two persons being malnourished if the trend continue. There are different forms of malnutrition namely undernutrition, micronutrient deficiency, overweight, obesity and diet-related non-communicable disease.

The Malawi nutritional profile (2018) explains that globally almost 800million people are undernourished (underweight, wasted and stunted) . Specifically, 155million under five children are stunted and 52million are wasted. In addition, report further explains that in low and middle income countries 45% of deaths in under five children are linked to under nutrition. In Malawi 37% of under five children are stunted, 3% are wasted and 12% are underweight according to the (2018) Malawi nutritional profile.

Micronutrient deficiencies is another form of malnutrition, it refers to inadequate intake of vitamins and minerals. Vitamins and minerals of great public health concern include vitamin A, iron and iodine. Lack of vitamin A increase the risk of diseases and death from infectious diseases and also cause blindness (Mphamba et al., 2024). Iron-deficient anaemia is a global concern for women since it affects cognition and work productivity. Iodine deficiency especially in pregnant women affects child's mental health. Other vitamins and minerals of importance include vitamin D, B12, folate, calcium and zinc (Nyanhanda et al., 2023).

Overweight, obesity and diet-related non-communicable diseases is another form on malnutrition that equally affects under five children. In 2014, 1.9billion people worldwide were overweight of which 600million were obese (G. N. Report, 2020). An estimated number of 41million under five children were overweight in 2014 and quarter of these children were from Africa. In Malawi currently 5% of under-five children are overweight. The high overweight rates are linked to increasing cases of diet-related non-communicable diseases such as diabetes, cardiovascular and cancer. Worldwide overweight is associated with more deaths than under-weight (Tesfay et al., 2022).

Globally, every country is affected by one or more forms of malnutrition. Pregnant and lactating women, infants, children, and adolescents are at particular risk of malnutrition. Improving nutrition early in life especially in the first 1000 days from child conception ensures the best possible start in life, with long-term benefits. On the other hand, poverty amplifies the risk of, and risks from, malnutrition. People who are poor are more likely to be affected by different forms of malnutrition (Mphamba et al., 2024). Furthermore, malnutrition increases health care costs, reduces productivity, and slows economic growth, which can perpetuate a cycle of poverty and ill-health ((WHO), 2020).

The need for countries to combat malnutrition in all its forms is one of the greatest global health challenges. In the global fight against malnutrition, in April 2016, the United Nations (UN) General Assembly proclaimed 2016–2025 the United Nations Decade of Action on Nutrition. The decade was set for countries to implement interventions to address all forms of malnutrition. The world nutrition association (2021) reports that UN set timeline for implementation of the commitments made at the Second International Conference on Nutrition (ICN2) to meet a set of global nutrition targets and diet-related NCD targets by 2025. Further, the UN news centre (2015) reports that the UN general assembly supports implementation of the

Agenda for Sustainable Development by 2030 thus Sustainable Development Goal (SDG) 2 , end hunger, achieve food security and improved nutrition and promote sustainable agriculture and SDG 3 ensure healthy lives and promote wellbeing for all at all ages . The decade is led by World Health Organization and Food and Agriculture organization. The UN Decade of Action on Nutrition calls for policy action across 6 key areas that includes creating sustainable, resilient food systems for healthy diets; providing social protection and nutrition-related education for all; aligning health systems to nutrition needs, and providing universal coverage of essential nutrition interventions; ensuring that trade and investment policies improve nutrition; building safe and supportive environments for nutrition at all ages; and strengthening and promoting nutrition governance and accountability, everywhere.

As a control measure for addressing all forms of malnutrition all policies and nutritional interventions in Africa and across the world are implemented in response to the UN Decade of Action on Nutrition. The global progress in combating malnutrition has been minimal. The (G. N. Report, 2020) reveals that in the 2018 assessment of progress against targets for nutrition only 94 countries out of 194 countries were on track in addressing malnutrition. Malawi was on track on 2 of the set 9 nutritional targets. Global stunting rate for under five children reduced from 32.6% in 2020 to 22.2% in 2017. There has been a slight decrease in underweight for women from 11.6% to 9.7% in 2016. In Malawi there has been a decrease in the malnutrition rates from 2000 to 2015-16. Stunting decreased from 55% to 37%, Underweight from 20% to 12%, wasting from 7% to 3% and overweight from 8% to 5% (Profile, 2018)

1.6 Nutritional Interventions in Malawi

Adequate nutrition is vital for human, physical and intellectual development (Machira & Chirwa, 2020). Malawi as a country has implemented different interventions to combat malnutrition. Amongst the interventions Malawi developed the National Multi-sectoral Nutritional Policy (2018-22). The policy was launched in 2018 with the aim of guiding the multi-sectoral nutrition response in terms of nutritional programming focusing on the national nutrition priorities which among others includes controlling acute malnutrition, preventing under nutrition and nutrition monitoring and evaluation research and surveillance. Also, the department of nutrition was established in (2018) to oversee implementation of nutritional programs . In addition, Malawi joined the scaling up nutrition movement in March 2011 and established a nutrition committee chaired by the secretary for nutrition, HIV and AIDS in the office of president. The movement's function is to mobilise resources towards implementation of interventions in line with the food and nutrition security policy (2005) and National policy and strategic plan (2015 -2017). The scaling up nutrition movement (2015) also focus on community based action to reduce stunting through behaviour change and awareness through the 1000 days of nutritional education and communication strategy for 2012 to 2017.

World Bank played a role in helping Malawi as a country to reduce cases of malnutrition in under-five children. In April 2016, Malawi benefited from a 26 million dollars grant of which more than half was used to combat the health problem of malnutrition in under five children caused by drought related food shortages (G. N. Report, 2021). The interventions implemented included community based capacity building for active case finding/screening, referral to nearby health centres, treatment of malnutrition cases and common illness associated with malnutrition. Care givers for children (6 to 36 months) with acute malnutrition were receiving therapeutic food chiponde (2%) and likuni phala (5%) as treatment. Chiponde is peanut butter based therapeutic food on the other hand Likuni phala fortified soya enriched flour for making porridge for malnourished children (Bank, 2016)

.

1.7 Problem statement

United Nations general assembly through the 2016–2025 Decade of Action on Nutrition lobby countries to implement interventions to address all forms of malnutrition which includes undernutrition by 2025 (Association), 2021). The fight agaisnt malnutrition cannot be succesful without a thorough review of dietary patterns since diet has a great contribution towards malnutrition (of Nutrition HIV & AIDS, 2018).

A number of studies have been done globally to explore the relationship between diet, undernutrition and other diseases affecting different groups of people. Silvera et al. (2011), Lin et al. (2007) and Batis et al. (2016) conducted studies to investigate the linkage between diets and undernutrition outcomes using statistical methods. Silvera et al. (2011) conducted a study on principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer. The study used PCA to itentify dietary patterns and logistic regression to assess the linkange between lifestyle patterns and Cancer. in addition, Lin et al. (2007) conducted a prospective assessment of food and nutrient intake in a population of Malawian children at risk for kwashiorkor , the study used regression modelling. Furthermore Batis et al. (2016) conucted a study using both principal component analysis and reduced rank regression to study dietary patterns and diabetes in Chinese adults. The study used PCA and RRR to derive dietary patterns and logistic regression to examine the relationship between dietary patters and diabetes.

In Malawi, Machira (2020) conducted a study to explore the dietary consumption and it's effect on nutrition outcomes among under five children in Malawi. The study used PCA to create factor variables and applied nested logistic regression to estimate the factors that impact the prevalence of undernutrition in Children. The study found that dietary consumption had an effect on underweight,stunting and wasting and recommended consistency in consumption of dietary foods to reduce nutritional deficiencies and sustainably deal with nutritional outcomes and its challenges.

Few studies have used statistical methods to link dietary patterns to undernutrition in underfive children in Malawi. Therefore, it is the purpose of this study to assess the linkage that exist between dietary patterns and nutritional outcomes in under five children in Malawi using Principal Component Regression (PCR), Reduced Rank Regression (RRR) and binary logistic regression.

1.8 Study objectives

1.8.1 Main objective

To model the link between dietary patterns and undernutrition outcomes in under five children in Malawi.

1.8.2 Specific objectives

1. Use principal component regression and reduced rank regression to derive dietary patterns

2. Apply logistic regression model to assess the linkage between dietary patterns and undernutrition outcomes

3. Investigate demographic factors associated with undernutrition outcomes.

1.9 Significance of the study

The results of the study will help to explore a statistical method which is relevant and useful in modeling the linkage between illnesses and their relevant cause in particular undernutrition and dietary patterns. In addition, the results will also provide information that would help in implementation of relevant interventions that contributes to improving nutritional status of under five children in Malawi and contributing to ending all forms of malnutrition including achieving the internationally agreed targets on stunting and wasting in under five children (Bank, 2016).

Additionally, the study results will contribute to the achievement of the Malawi multi-sector nutrition policy 2018-2022 priority areas. These include priority area number 1 on preventing undernutrition and priority area number 2 on ”gender equality , equity, protection, particiption and empowerment for improved nutrition” as specified by the department of nutrition on HIV and AIDS (2018). The results will also contribute to the fullfillment of priority area number 6 in the Malawi: Poverty Reduction Strategy Paper—Growth and Development Strategy (2007)thus Prevention and Management of Nutrition Disorders, HIV and AIDS.

Furthermore, it will contribute to the achievement of sustainable development goals. These include SDG 1 ”end poverty in all its forms everywhere”, SDG 2 ”end hunger,achieve food security and improved nutrition and promote sustainable agriculture”, SDG 3 ”ensure healthy lives and promote well being for all at all ages”, SDG 4 ”ensure inclusive and equitable quality education and promote life-long learning opportunities for all”,and SDG 8 ”Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all” By 2030 (UN news centre) (2015).

1.10 Thesis structure

The thesis is structured as follows, Chapter two describes literature and data overview including basic descriptive analytics of the data. Chapter three presents principal component regression on dietary patterns. Chapter four presents reduced rank regression on dietary patterns and chapter five describes the overall discussion, conclusion and recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter describes principal component regression method, reduced rank regression method , logistic regression method and their application in assessing the linkage between dietary patterns and it's outcomes.

2.2 Principal Component Regression Method

Principal component regression is a statistical method that uses principal component analysis to summarize original predictor variables into a small set of variables called principal components which are a linear combination of the original predictor variables (Mukherjee, 2013) . The selected subset of the principal components are used to build the regression model, thus the principal components are used as regressors. PCR has three main steps, the first one is obtaining principal components and selecting subset of the principal components to be included in the analysis based on set criteria. Secondly, regressing the vector of outcomes on the selected subset of principal components as predictor variables using ordinary least squares (OLS) regression to get coefficients and the last step is transforming the vector back to the scale of the actual predictors, using the selected PCA loadings to get the final PCR estimator for estimating the regression coefficients characterizing the original model.

Selection of principal components to be included in principal component analysis depends on the eigen value from a particular eigen vector. Eigen vectors reflect both common and unique variance of the variables. Kaiser criterion recommends that principal components with eigen values greater than one should be retained and included in analysis (Braeken & Assen, 2017).

Eigen values measure the variance in all the variables which is accounted for by that factor. This helps to produce a low dimensional representation of a data set. The aim is to find a linear combination of variables with maximum variance and mutually correlated. Scree plots are also used to cross-validate appropriate number of principal components to be retained. Scree plots orders the eigen values from largest to smallest, the cut off point is where there in as "elbow" and principle components before the elbow are included in the analysis. The first principal component accounts for the most variation as shown in figure 1 .

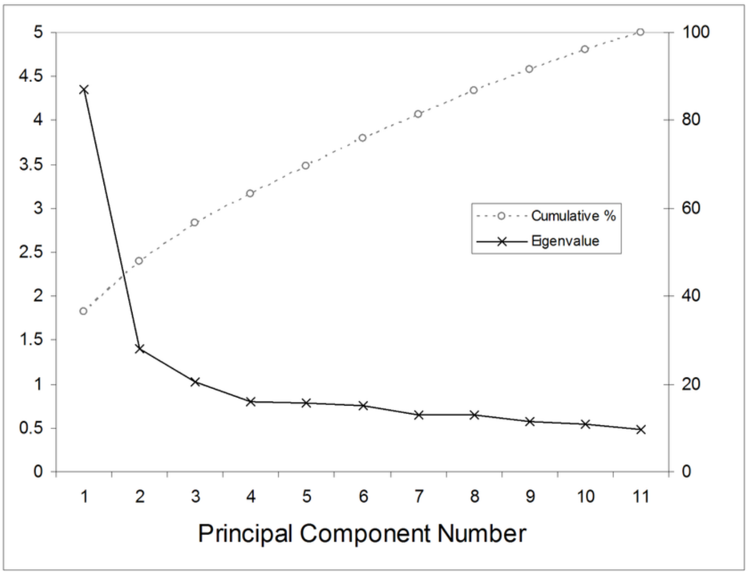

Figure 1: Principal components Analysis Scree Plot and Cummulative Variance Plot

Output from PCA contains set of eigen values that provides variability information in the data, table with principal components from which a set of components are selected and table of loadings describing correlation between variables and principal components that helps getting the relationship between variables as well as their associations with the extracted principal components.

Suppose we have a vector of outcomes

$$\mathbf{Y}_{n\times 1} = (y_1, \cdots y_n)^T \tag{2.1}$$

denote the vector of oberved outcomes and

$$\mathbf{X}_{n\times p} = (x_1, \cdots x_n)^T \tag{2.2}$$

denote data matrix of oberved predictors where n denotes the size of the observed sample and p denotes the number of predictors, $n \geq p$, each of n rows of $\mathbf{X}$ denotes one set of observations for the p dimensional predictors and each $\mathbf{Y}$ entry denotes the corresponding observed outcome (Kramer, 1998).

The first principal component has the form

$$C1 = b_{11(X_1)} + b_{12(X_2)} + \cdots + b_{1p(X_p)} \tag{2.3}$$

where b_{1p} is the regression coefficient X_p is the subject's score on observed variable p and for the loadings equation

$$Z_1 = \phi_{11}X_1 + \phi_{21}X_2 + \cdots + \phi_{p1}X_p \tag{2.4}$$

elements $\phi_{11} \cdots \phi_{p1}$ are loadings of the first principal component which is given by the loadings vector

$$\phi_1 = \phi_{11}, \phi_{21}, \cdots, \phi_{p1}^T \tag{2.5}$$

Sanchez and Marzban (2020) explains that in PCR we seek to find principal components which are linear combination of inputs hence the standard Gauss Markov linear regression model for PCR is given by

$$Y = X\beta + \varepsilon \tag{2.6}$$

where $\beta \in \Re^p$ denotes the unknown parameter of regression coefficient and ε denotes the vector random error with $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 I_{n\times n}, \sigma^2 > 0$

Regression is performed on the assumption that matrix **Y** of the observed outcomes and each of the p columns of **X** have already been centred having mean zero.

The objective is to get an efficient estimator $\hat{\beta}$ for the parameter β based on the observed data using ordinary least squares regression.

The main assumption in PCR is that the direction in which predictors show the most variation is the exact direction associated with the response variable.

$$\hat{\beta} = (X^T X)^{-1} X^T Y) of \beta \tag{2.7}$$

2.2.1 Justification of Using principal Component Regression

Statistical analyses of linking diets or foods to undernutrition involves multiple variables, there are multiple variables in this study that require dimension reduction methods hence the need for principal component analysis. Principal component regression is a dimension reduction tool used to reduce a large set of predictor variables into a small set of variables containing most of the information to be used in a model (Kramer, 1998).

PCR helps reduce dimensionality and fit regression model to a smaller set of variables. It converts a set of correlated observations into a set of linearly uncorrelted variables. PCR is useful in reducing multi-dimensional data to lower dimensions whilst keeping most of the information, thus it reduces model complexity which lead to efficient prediction of the built model. PCR also reduces multicolinearity since using PCA on the raw data produces linear combinations of the predictors that are uncorrelated. It is suitable when the data set contains highly correlated predictors, thus it helps to overcome the problem of multicolinearity by excluding some of the low-variance principal components in the regression (Ali, Margetts, & Zainuddin, 2021).

2.3 Reduced Rank Regression Method

Reduced rank regression model is an estimation method in multivariate regression model. It uses reduced rank algorithm to estimate the coefficient matrix with reduced rank. RRR takes into account the reduced rank restriction on the coefficient matrix. (Zheng, Liu, Lyu, & Yu, 2022) It involves the multivariate regression of response variable Y of dimension p on matrix X and Z of dimension q and k respectively. The model is given by the following equation.

$$Y_t = \Pi X_t + \mathit{\Gamma} Z_t + \varepsilon_t \tag{2.8}$$

where $t = 1, \cdots, T$

$\Pi = \alpha\beta'$ is the hypothesis that Π has reduced rank less than or equal to r where α is $p \times r$ and β is $q \times r$ where $r < min(p, q)$ which gives a reduced rank model

$$Y_t = \alpha\beta' X_t + \mathit{\Gamma} Z_t + \varepsilon_t \tag{2.9}$$

where $t = 1, \cdots, T$

Reduced rank regression algorithm is given by $RRR(Y, X/Z)$; it involves 3 steps

1. Regressing Y on X and Z and form residuals

$$(Y/Z)_t = Y_t - S_{yz}S_{zz}^{-1}Z_t \tag{2.10}$$

and

$$(X/Z)_t = X_t - S_{xz}S_{zz}^{-1}Z_t \tag{2.11}$$

and the product moments

$$S_{yx.z} = T^t \sum \min t = 1 \max T(X/Z)_t((X/Z)_t)^1 = (S)_{yx} - S_{yz}{S_{zz}}^{-1}Z_t \tag{2.12}$$

It should be noted that reduced rank regression of Y and X is an algorithm for estimating the parameter of interest β using matrix X

2. Solving the eigen value problem

$$|\lambda S_{xx.z} - (S)_{xy.z}{S_{zz}}^{-1}S_{yx.z}| = 0 \tag{2.13}$$

where the right side denotes determinant. The ordered eigen values are $\Lambda\Lambda = diag(\lambda_1, \cdots, \lambda_q)$ and the eigen vectors are $V = (v_1, \cdots, v_q)$ so that

$$S_{xx.z}V\Lambda = (S)_{xy.z}{S_{yy.z}}^{-1}S_{yx.z} \tag{2.14}$$

3. Defining estimators

$$\hat{\Pi}_{RRR} = S_{yx \cdot z}\hat{\beta}(\beta^1 S_{xx.z}\hat{\beta})^{-1}\hat{\beta}^1 \tag{2.15}$$

of the coefficient matrix to X

multivariate vectors $Y \in \Re^N$ and $X \in \Re^N$ satisfy $E(X) = 0, E(Y) = 0, E(Y/X) = BX$

2.3.1 Justification of Using Reduced Rank Regression

Reduced rank regression is a statistical method used in analyses that involve a number of covariates. It is widely used in nutritional epidemiology in modelling dietary patterns and their outcomes. Dietary patterns data contains many covariates hence it is ideal to use RRR. This method helps to obtain few predictors from a large coefficient matrix. It is an effective method in predicting multiple response variables from set of predictors. RRR helps prediction accuracy and facilitates parameter estimation and model intepretation. (Mukherjee, 2013)

2.4 Logistic Regression Method

Logistic regression model is a multiple regression method that is used to analyse the relationship between binary outcome or categorical outcome with a number of explanatory variables (Soetewey, 2024). Types of logistic regression model include binary logistic model, multinomial logistic regression model and ordinal logistic regression model. Binary logistic regression is applied when the dependent variable has two outcomes is yes or no for instance presence of a disease or not.

Multivariate logistic regression is used when there are multiple predictors (ref Nick). Ordinal logistic regression is used when a response variable is categorical with clear ordering of the category levels (Parry, 2020).

The study investigated the linkage between dietary patterns and undernutrition outcomes thus stunting, wasting and underweight. The study used binary logistic regression because it aimed at examining the linkage between dietary patterns with each of the undernutrition outcome. Binary logistic model produces odds ratios which suggest increased, decreased or no change in odds of being in one category of the outcome with an increase in the value of the predictor. Model significance quantifies whether the model is better than the baseline value at explaining whether the observed cases in the data set have the outcome (Harris, 2021).

Binary logistic regression relies on three underlying assumptions to be true. The option include the observations must be independent, there must be no perfect multicollinearity among independent variables and continuous predictors are linearly related to a transformed version of the outcome Soetewey (2024) defines the binary logistic regression equation as presented below

$$p\left(y\right) = \frac{1}{1 + e^{-(b_0 + b_1 x_1 + b_2 x_2)}}$$

In addition Winter (2024) describes Logit link used in the study is presented by below

Logit $$\log\left(\frac{\pi_i}{1-\pi_i}\right)=\mathbf{x}_i^T\boldsymbol{\beta}$$

Odds and odds ratio (OR) of Binary logistic regression models are used to interpret logistic regression results. The estimate b0 and b1 as specified in the binary logistic regresssion model are estimated using statistical software (Fariq & Fatah, 2016). After the values are estimated they are substituted into the equation for computing odds and odds ratio .Odds ratio is computed by dividing the odds of the outcome at one value of a predictor by the odds of the outcome at the previous value(Soetewey, 2024). The odds and odds ratio equations are presented below

$$odds=\frac{\frac{1}{1+e^{-(\beta_0+\beta_1 x)}}}{1+\frac{1}{1+e^{-(\beta_0+\beta_1 x)}}}=e^{\beta_0+\beta_1 x}$$

$$OR=\frac{e^{b_0+b_1(x+1)}}{e^{b_0+b_1 x}}=e^{b_1}$$

2.5 Sample Studies that Applied Principal Component Regression , Reduced Rank Regression and Logistic Regression on Dietary Patterns Analysis and its Outcomes

2.5.1 Application of principal component regression and Logistic Regression in dietary analysis

Studies have been done to assess the linkage between dietary consumption and nutritional outcomes worldwide. In Ethiopia, a study was conducted by (Kassie & Workie, 2020) using principal component analysis to assess determinants of under-nutrition in underfive children. The study used demographic and health survey data collected in Ethiopia in 2016 with a sample of 9494 children below 59 months. The survey found out that under-nutrition is associated with different factors including husband's education status, mother's body index, number of under-five children and wealth index among others. The authors recommended prioritizing interventions that focus on household food security, wealth index, mothers/spouses education, increasing mother's health care access inorder to combat under-nutrition in under five children.

Furthermore, another study done by (Moradi, Jalilpiran, Askari, Surkan, & Azadbakht, 2022) used principal component factor analysis to investigate the associations between mother and child dyad dietary patterns and child anthropometric measures among 6 year old children. This study used data from a cross sectional study conducted in health centers located in Tehran, Iran. A total of 788 pairs of 6 year olds and mothers were recruited in the study and dietary consumption information was collected using a food frequency questionnaire with 168 items. Dietary patterns were derived using principal component factor analysis and associations were evaluated using binary logistic regression model. The study found that there was an inverse association between high protein dietary patterns in children and the undernutrition outcome of being underweight and wasted.

Moreover, a more outcome specific study was done in Nepal by (Shrestha, Shrestha, & Cissé, 2021) where principal component analysis was used to assess the association of dietary patterns and stunting among Nepalese school children. The study used data from 708 school going children from 8 to 16 years old, collected in a cross-section baseline study in the two districts of Dolakha and Ramechhap. Interviews were conducted with caregivers where they provided data on food intake. Dietary patterns were derived using principal component analysis and the linkage on stunting was analysed using mixed logistic regression. The results showed that consumption of vegetable and animal protein is associated with reduced odds of stunting. All the listed studies shows that there is a link between dietary patterns and undernutrition outcomes.

(Ali et al., 2021) used Principal Component Analysis to examine dietary patterns of the Malaysian population. The study sampled 3063 respondents aged 18 to 59 years from Peninsular Malaysia, Sabah, and Sarawak who reponded to the food frequency questionnaire in the Malaysian Adult Nutrition Survey which was conducted from 2002 to 2003. Correlation between dietary patterns were derived using Principal Component Analysis and selected nutrients intake were determined. With Kaiser criterion the results of the survey showed that six components thus dietary patterns namely traditional, prudent, modern, western, chinese and combination diets with eigen value greater than 1 explained 45.9% of the total variability.

In addition, the components derived from factor loadings showed positive association with several nutrient markers. Traditional dietary pattern revealed a moderate positive correlation with total protein and total sugar intake. The Prudent dietary pattern showed a significant moderate correlation with dietary fibre, and a positive association between the Chinese dietary pattern and total energy was also shown. (Ali et al., 2021) reccomended that the PCA performed provides a justification for thorough assessment of dietary patterns instead of relying on single nutrients or foods in identification of potential connections to overall nutritional wellbeing and also in exploring the diet–disease relationship.

Another study done by (Zhen, Ma, Zhao, Yang, & Wen, 2018) evaluated the association of dietary patterns with obesity in chinese children and adolescents using Principal Component Method. The study used data from a longitudinal China Health and Nutrition Survey (CHNS) where 736 children between ages 6 – 14 were followed from 2006 to 2011. Principal Component Method was used to determine the dietary patterns by factor analysis during baseline. The dietary patterns were identified on the basis of the eigenvalue that was greater than 2, scree plots were ploted and percentage of variances were calculated. Factor loading of $> |0.20|$ was used to represent the food strongly associated with the identified factor. Factor scores were divided into four quartiles on the basis of their contribution to each pattern,to compare the different quartiles in dietary patterns ANOVA tests for continuous variables and chi-square tests for categorical variables were used at baseline.

The survey revealed two dietary patterns this included traditional chinese dietary pattern and the modern dietary pattern. The survey results show an association between dieatry pattern and the nutritional outcome that is obesity. Children in the highest quartile and the second highest quartile consuming the traditional chinese dietary pattern were obese compared with children in the lowest quartile over the 5 years. On the other hand "children in the highest quartile of the modern dietary pattern were positively associated with later obesity compared with children in the lowest quartile over 5 years".

Furthermore, a study done by (Batis et al., 2016) on "using both principal component analysis and reduced rank regression to study dietary patterns and diabetes in Chinese adults". The study used principal component analysis and reduced rank regression to examine the association between dietary patterns and diabetes in chinese adults. PCA was used to determine the eating patterns of the population on the other hand RRR was used to derive a pattern that explains the variation in glycated Hb (HbA1c), homeostasis model assessment of insulin resistance (HOMA-IR) and fasting glucose. The study used a sample of 4316 adults from the China Health and Nutrition Survey. Over 3 days diet with 24 hours recall and a household food inventory in 2006 was used to derive PCA and RRR dietary patterns.

Dietary patterns were analysed using twenty-nine food groups. PCA was performed using the procedure PROC PLS with a PCR method option and a total of ten components had an eigenvalue greater than 1. However there was a clear break in the scree plot after the second component and only the first two components were retained. Furthermore, RRR was performed with PROC PLS and an RRR option with HbA1c, HOMA-IR and fasting glucose as response variables. Natural log transformation of all response variables was done due to non-normality . RRR analysis only returned one dietary pattern since it was the only one with significant association and explained most of the variation in the response variables. Scores were calculated for the dietary patterns as weighted sum of the food groups based on the factor loadings. The higher the score, the more closely the participant's diet conformed to the dietary pattern. multiple linear (for HbA1c, HOMA-IR and fasting glucose) and logistic (for diabetes) regressions was performed for each dietary pattern to identify how different the strength of association was between RRR and PCA factors.

The analysis showed that for the adjusted odds ratio for diabetes prevalence ($HbA1c >= 6 \cdot 5\%$), comparing the highest dietary pattern score quartile with the lowest, was 1.26 (95 % CI 0·76, 2·08) for a modern high-wheat pattern (PCA; wheat products, fruits, eggs, milk, instant noodles and frozen dumplings), 0·76 (95 % CI 0·49, 1·17) for a traditional southern pattern (PCA; rice, meat, poultry and fish) and 2·37 (95 % CI 1·56, 3·60) for the pattern derived with RRR. Comparison of the dietary pattern structures derived using PCA and RRR showed that patterns derived from RRR were behaviorally meaningful. The study showed that using PCA and RRR provides insights in studying the association of dietary patterns with diabetes

2.5.2 Application of reduced rank regression and Logistic Regression in dietary analysis

Studying association of dietary patterns with nutritional outcomes and other illnesses using reduced rank regression has been done by researchers in nutritional epidemiology. (Curtis et al., 2023) used reduced rank regression in a study to examine the association of dietary patterns and malnutrition , muscle loss and sarcopenia in cancer survivors. The study used data from a cross-sectional study that involved 2415 cancer survivors from from UK Biobank. Dietary patterns were derived using RRR and logistic regression analysis was performed to examine the association. The results showed energy rich dietary patterns were associated with lower odds of being malnourished. Specifically, meat and dairy dietary pattern were not associated with malnutrition on the other hand cancer survivors who were consuming highly oily fish and nuts had significantly lower odds of being malnourished.

In another study done by (Huybrechts et al., 2017), reduced rank regression was used to identify dietary patterns associated with obesity. This was a cross sectional study among European and Australian adolescents. The study used data from two cross sectional surveys in Europe with 1954 adolescents and Australia 1498 adolescents. 24 hour recalls on dietary intake was measured. Dietary patterns were derived using RRR and association was examined using multivariate linear and logistic regression. The results showed that "energy dense, high fat , low fibre" dietary patterns explained 47 % and 31% of obesity in Australian and European adolescents. This shows that dietary patterns are associated with nutritional outcomes.

Other studies have done factor specific investigation on dietary pattern and nutritional outcomes. For instance, (Mayasari et al., 2023) examined the relationship between dietary patterns and erythropoiesis associated micronutrient deficiencies (iron, folate and vitamin B_{12}) among pregnant women in Taiwan. The study aimed at identifying dietary patterns for preventing erythropoiesis associated micronutrient deficiencies in pregnant women. Diet, anthropometrics and biochemistry data was collected from 1437 pregnant women aged between 20 and 48 during prenatal visits. RRR was used to identify dietary patterns. The study found that there was a positive correlation between the dietary pattern (nuts and seeds, fresh fruits, vegetables, breakfast cereals/oats, soybean and dairy products) and micronutrient deficiencies in anemic pregnant women. This shows that dietary patterns are associated with nutrient deficiencies.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the methods used in this study. It gives an overview of sampling, data source and data analysis.

3.2 Sampling

The study used data collected during the cross-sectional study of the 2015-16 Malawi Demographic and Health Survey (MDHS) by National Statistical Office (NSO). MDHS used the 2008 Malawi population and Housing Census (MPHC) frame to obtain sample for the survey. This frame is a complete list of all census standard enumeration areas (SEA) which was created for the 2008 census. An enumeration area is described as a geographical area that covers an average of 235 households. The data was collected from 28 districts in Malawi. Sampling of households was done in two stages, the first stage used stratified sampling to select SEAs which were also taken as clusters. A total of of 850 SEAs 677 in rural areas and 173 in urban areas were selected. Household listing was done in all the SEAs which served as a sampling frame for selection of households in the second stage of sampling. A fixed number of 30 households in urban clusters and 33 households in rural clusters were selected with equal probability systematic to form the new household list. In total 25,146 households were sampled thus 5,363 in urban and 19, 783 in rural areas respectively. Data on the foods and liquids consumed by the under –five children and anthropometry data for the under-five children were collected from 6,033 households thus a subsample of one third of the households and nutrition status of 4151 children was reported by their mothers (Malawi-NSO & ICF, 2017) .

3.3 Data Source

Feeding practices and anthropometric data that measures nutritional status of under 5 Children was used in the study as captured during the cross-sectional study of the 2015-16 Malawi Demographic and Health Survey (MDHS) by National Statistical Office (NSO). NSO captured data on child health nutrition in the woman's questionnaire where a mother or caretaker was asked a list of foods and liquids that each of the under five child in the house consumed 24 hours before the interview (Malawi-NSO & ICF, 2017). This was done considering that diversity of foods and frequency of feeding have effect on nutritional status of a child. The MDHS report outlines that generally, no major variations exist on consumption of complementary foods between breastfed and non-breast fed infants hence the analysis focused on all under-five children regardless whether the child was breastfed or not (Malawi-NSO & ICF, 2017).

On the other hand, anthropometry data that includes weight and height was collected using the biomarker questionnaire. Height of the children was measured using a Shorr Board. Children under 2 years of age were measured lying down on the board whilst the rest were measured standing. Furthermore, Weight of the children was measured with an electronic SECA 878 flat scale. Weight for young children was measured by subtracting weight of the mother or caregiver from the weight of the mother or caregiver whilst holding the child. The weight and height data was used to calculate height-for-age (stunting), weight-for-age (underweight) and weight-for-height (wasting). Anthropometric data allows measurement and evaluation of nutritional status in children. It helps to identify subgroups of child population at risk of waned growth, disease, impaired development and death (Machira & Chirwa, 2020).

The WHO standard definition of stunting, underweight and wasting was used in the study. The definition explains that z-score for weight-for-age (WAZ) below -2 standard deviation (SD) describes the child as underweight. (Mphamba et al., 2024) Z score for height-for-age (HAZ) below -2SD is described as stunting and defines wasting as weight-for-height (WHZ) below -2SD. Under-nutrition measures were defined as a binary variable in which if the Z score of WAZ, HAZ and WHZ was less than -2SD, the recoded value was 1 which meant the child in under-nourished (underweight, stunting and wasting) respectively, and 0 was otherwise (Maleta, 2006).

3.4 Data Analysis

Data analysis was done using R software version 3.6.1 (2019-07-05). However before analysis the data was cleaned in Stata software version 14.0. Missing values and outliers were cleaned from the data set. Data cleaning yielded a list of 17 food and liquids which children consumed day and night before the survey. Only food and liquids consumed by 5% of the under five chidren were considered to be used in the analysis as used in a similar previous study by (Batis et al., 2016). The data was precoded in three categories of Yes, No and I dont know. These 17 food items were used as covariates and 3 nutritional outcomes namely stunting, wasting and underweight, were used as outcome variables in the analysis.

Descriptive analysis was performed to show foods and liquids consumption and demographics of the children included in the analysis. Principal Component Analysis was performed in R software version 3.6.1 using the function Princomp to get principal components. The outcome yielded 17 components of dietary patterns. Further, Kaiser Criterion was used to obtain eigen values. Three components had an eigen value greater than 1. The eigen value represents a partitioning of the total variation accounted for by each principal component.

A scree test was done to determine statistically significant components to be retained. The procedure uses a scree plot. The cut off point is where there is a break with eigen values greater than 1. After ploting the scree plot, the first 3 components were returned. Loadings of the principle components were also computed in R. The loadings helps to determine foods and liquids to be included in the models, those with principal loadings $\geqslant |0.25|$ were included in the models. Logistic regression for underweight, stunting and wasting were performed with the retained principal components and with control variables to get the models.

Reduced rank regression of type "pca" was performed in R to get dietary patterns. Loadings of the reduced rank regression were also run, significant components of dietary patterns were derived and a goodness of fit was performed on the components.

3.4.1 Descriptive Analytics of the Data

The list of foods and liquids included in the analysis is shown in Table 1. A total sample of 9103 children was included in the analysis of which 49 % were male and 51 % were female. The dietary consumption percentage data in Table 2 presents food consumption presentage by type of residence. The data shows that more of the food items were consumed by children living in the urban areas than rural areas

Table 1: Percentage of foods and liquids consumption for children in the day or night before the interview.

Food Group	Standard Deviation (SD)	Consumption Percentage	Overal Group Consumption Percentage
Carbohydrates			
Bread, noodles, other food made from grains	0.53	48%	42%
Other solid and semi-solid food	0.59	44%	
Soup/Clear broth	0.53	43%	
Potatoes, cassava and other tubers	0.40	10%	
Fortified baby food (cerelac)	0.30	6%	
Legumes and Nuts			
Beans, Peas, lentils, nuts	0.53	18%	5%
Fruits			
Mangoes, pawpaw and other vitamin A fruits	0.55	34%	15%
Any other fruit	0.45	21%	
Vegetables			
Dark green leafy vegetables	0.58	45%	12%
Animal food			
Fish, shell fish	0.44	17%	9%
Eggs	0.39	9%	
Meat (beef, pork, lamb, chicken)	0.39	8%	
Fats			
Oil, fats, butter and products made from them	0.37	9%	3%
Other Foods and Drinks			
Other liquid	0.50	24%	14%
Juice	0.36	11%	
Chocolate, sweets, candies and pastries	0.45	10%	
Soft drinks	0.35	5%	

Table 2: Percentage of foods and liquids consumption for children in the day or night before the interview by type of place of residence.

Food Group	Consumption percentage by type of place of residence	
	Urban	**Rural**
Carbohydrates		
Bread, noodles, other food made from grains	52%	48%
Other solid and semi-solid food	48%	44%
Soup/Clear broth	45%	44%
Potatoes, cassava and other tubers	10%	10%
Fortified baby food (cerelac)	14%	6%
Legumes and Nuts		
Beans, Peas, lentils, nuts	20%	18%
Fruits		
Mangoes, pawpaw and other vitamin A fruits	33%	35%
Any other fruit	24%	21%
Vegetables		
Dark green leafy vegetables	46%	45%
Animal food		
Fish, shell fish	20%	17%
Eggs	13%	8%
Meat (beef, pork, lamb, chicken)	13%	7%
Fats		
Oil, fats, butter and products made from them	13%	9%
Other Foods and Drinks		
Other liquid	29%	24 %
Juice	17%	10%
Chocolate, sweets, candies and pastries	13%	10%
Soft drinks	12%	4%

Food group consumption scores were calculated using WHO recommended weights for food groups. The scores were categorised using consumption threshold of 28/42 (Wiesmann, Bassett, Benson, & Hoddinott, 2009). The consumption rates namely acceptable, borderline and poor were calculated. Results showed that over 90% of the food groups had poor consumption rate.

The food groups consumed by the underfive children were categorised into 3 dietary diversity levels. The food groups were firstly categorised into 12 food groups using FANTA approach (Bilinsky & Swindale, 2006) and dietary diversity levels were derived. The dietary diversity levels include low dietary diversity, medium dietary diversity and high dietary diversity. The study applied household dietary diversity level definition used in a dietary diversity study by (Sambo, Oguttu, & Mbombo-Dweba, 2022). The levels are shown in 3

Table 3: Dietary Diversity Levels for underfive children

Dietary Diversity Level	**Range**	%
Low	1 - 3	57
Medium	4 - 5	26
High	6 - 12	17
Total	**12**	**100**

The trends of undernutrition outcomes of the underfive children in Malawi from 1992 to 2015-16 are shown in figure 2. The levels of stunting, wasting and underweight decreased from 1992 to 2015-16 with a more significant decrease in stunting from 2010 to 2015-16.

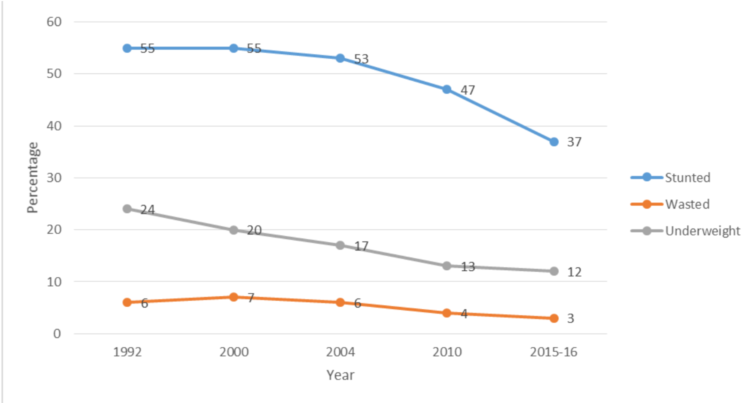

Figure 2: Trends in undernutrition status of underfive children in Malawi

On the other hand regional distribution of undernutrition outcomes shows that central region of Malawi has the highest stunting levels (37%) whilst the southern region has the highest wasting (4%) and underweight (13%) levels in Malawi 3.

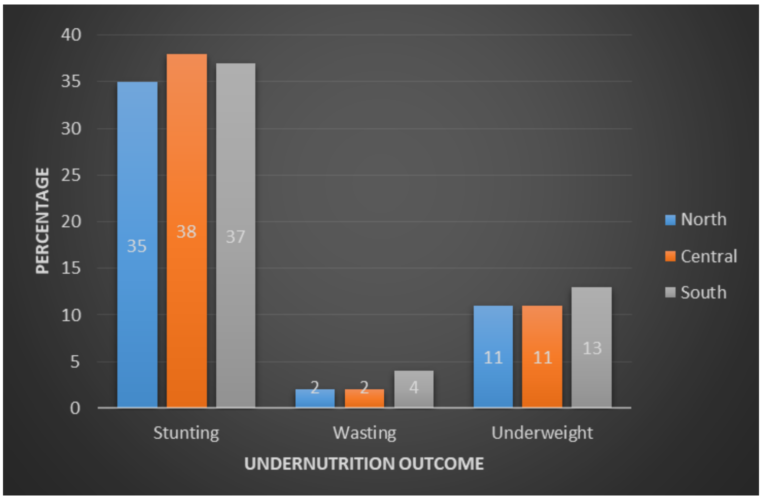

Figure 3: Regional distribution of undernutrition outcomes for underfive Children in Malawi

Moreover, undernutrition levels change with age of the child. Figure 4 shows that underweight levels in underfive children in Malawi increase as the age of the child increases. Stunting also increases between ages 6 to 8 months and 36 to 47 months.

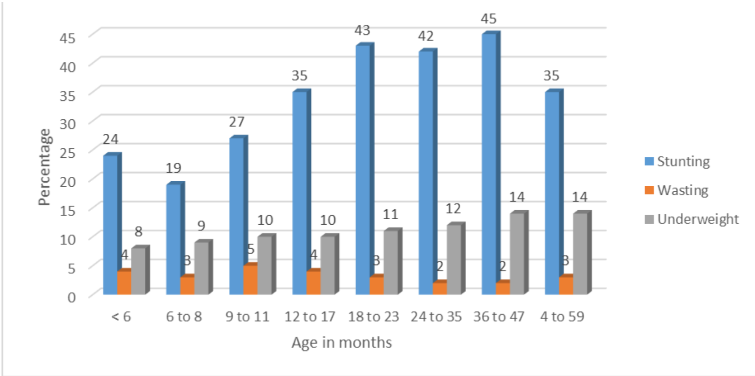

Figure 4: Age distribution of undernutrition outcomes for underfive Children in Malawi

CHAPTER 4

RESULTS

4.1 Introduction

This chapter provides an overview of the results of dietary patterns derived from PCR and RRR. It also provides logistic regression results of the linkage between the derived dietary patterns and undernutritional outcomes. The chapter finally describes other factors associated with undernutrition

4.2 Application of principal component regression and logistic regression to assess the linkage between dietary patterns and undernutritional outcomes

4.2.1 Principal component regression derived dietary patterns

The analysis outcome yielded 17 principal components. Using the Kaiser criterion where principal components with eigen values greater than one are retained, principal component 1 with eigen value 4 was retained.

The eigen value represents a partitioning of the total variation accounted for by the principal component. The proportion of variance explains the proportion of variance explained by the principal component. Principal component 1 explains 25% of the variation. A scree plot 5 was plotted as a useful visual aid for determining appropriate number of principal components to be retained. The scree plot graphs the eigen value against the component number.

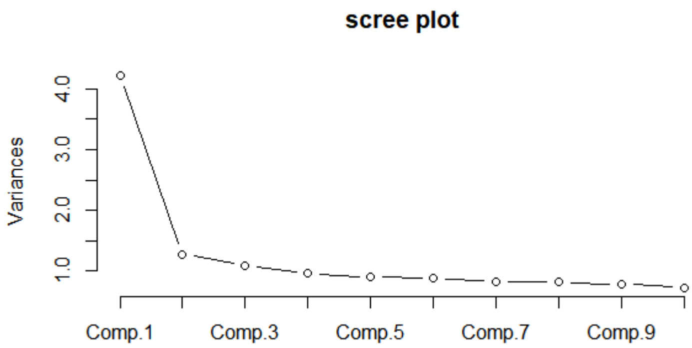

Figure 5: PCA eigen values scree plot

The scree plot qualifies components 1 to be retained in the analysis. The cut off point is where there in as "elbow" in the scree plot that's the point at which the remaining eigen values are relatively small and all about the same size.

Table 4 show the loadings of food groups for the retained principal component. The larger the absolute value of the coefficient, the more important the corresponding variable is in calculating the component. Food groups with loadings of $\geqslant |0.20|$ were included in analysis as used in a previous similar study done by (Zhen et al., 2018). Therefore, a total of 15 food groups with loadings with $\geqslant |0.20|$ were used to explain stunting wasting and underweight as listed in table 4.

Table 4: Loadings of the principal component

Food Group	Loadings
Cereals	
Bread, noodles, other food made from grains	0.246
Fortified baby food (cerelac)	0.208
Roots and tubers	
Potatoes, cassava and other tubers	0.245
Vegetables	
Dark green leafy vegetables	0.294
Fruits	
Mangoes, pawpaw and other vitamin A fruits	0.28
Other fruits	0.301
Meat, poultry, offal	
Meat (beef, pork, lamb, chicken)	0.232
Eggs	
Eggs	0.239
Fish and sea food	
Fish, shell fish	0.243
Pulses, legumes and nuts	
Beans, Peas, lentils, nuts	0.225
Milk and milk products	
Other liquid	0.246
Oil and fats	
Oil, fats, butter and products made from them	0.265
Sugar, honey	
Chocolate, sweets, candies and pastries	0.243
Miscellaneous	
Soup/Clear broth	0.248
Other solid and semi-solid food	0.27

The principal component regression derived dietary pattern is a carbohydrates-legumes-fruits-vegetables-animal-protein-fats-sugar pattern that will be used as an independent variable in the analysis.

4.2.2 Logistic regression modelling for PCR derived dietary patterns

The carbohydrates-legumes-fruits-vegetables-animal protein-fats-sugar dietary pattern derived from PCR was split into 3 categories which represented dietary diversity levels namely low, medium and high. This section outline results of the logistic regression model for the dietary diversity levels as independent variable with stunting, underweight and wasting as dependent variables. The section also gives results of fitting stunting wasting and underweight with control variables to investigate other factors associated with stunting, underweight and wasting.

4.2.2.1 Stunting binary logistic regression models for PCR dietary pattern

Table 5: Logistic regression model for stunting

	Estimate	**z value**	$Pr(>\|z\|)$
(Intercept)	-1.199	-16.281	$2e-16$
Dietary diversity score			
Low (ref)			
Medium	0.394	3.926	8.65E-05
High	0.261	2.555	0.010

AIC 3555.1

The result in table 5 show that at 5% significance level dietary diversity levels for the carbohydrates-legumes-fruits-vegetables-animal protein-fats-sugar pattern derived from PCR are significantly associated with stunting. This implies that children who were not consuming the carbohydrates-legumes-fruits-vegetables-animal protein-fats-sugar pattern in the medium (OR=1.483, p$<$ 0.05, cI: 1.218-1.806) and high (OR=1.298, p$<$ 0.05, CI:1.063-1.586) dietary diversity levels were likely to be stunted. This agrees with (Machira & Chirwa, 2020) who found that inadequate consumption of fruits, vegetables and carbohydrates was associated with stunting in underfive children in rural Malawi.

Table 6: Logistic regression model for stunting with demographic factors

	Estimate	**z value**	**Pr(> |z|)**
(Intercept)	-1.356	-6.279	$3.40E-01$
Dietary diversity level			
Low (ref)			
Medium	0.373	3.676	0.000
High	0.318	3.059	0.002
Sex of household head			
Male (ref)			
Female	0.070	0.739	0.460
Region			
North (ref)			
Central	0.355	2.692	0.008
South	0.277	2.174	0.030
Place of residence			
Urban (ref)			
Rural	0.334	2.564	0.010
Highest Level of education			
No education (ref)			
Primary	-0.190	-1.522	0.128
Secondary	-0.490	-3.008	0.003
Tetiary	-0.820	-1.747	0.081
Wealth Index			
Poorest (ref)			
Poorer	-0.048	-0.419	0.675
Middle	-0.253	-2.036	0.042
Richer	-0.442	-3.236	0.001
Richest	-0.469	-3.014	0.003

AIC: 3506.2

Fitting the stunting model with demographic factors , the output in table 6 show that place of residence , parent or guardian highest level of education, wealth index and region are associated with stunting since atleast one level of these factors is significant at 5% significance level ($p < 0.05$) (Soetewey, 2024). Specifically, children residing in the central and southern region of Malawi were likely to be stunted than those residing in the northern region of Malawi. This implies that place of residence has an effect in dietary consumption and nutritional status (Ntenda & Chuang, 2018). The full model fitted with dietary diversity levels and demographic factors is a better model with AIC 3555.1 compared to the stunting model fitted with dietary diversity levels only.

4.2.2.2 Wasting binary logistic regression models for PCR dietary pattern

Table 7: Logistic regression model for wasting

	Estimate	**z value**	$Pr(>\|z\|)$
(Intercept)	-2.960	-20.206	$< 2e-16$
Dietary diversity level			
Low (ref)			
Medium	-0.405	-1.761	0.078
High	-0.139	-0.647	0.518

AIC: 1032.1

The results for wasting logistic regression for this study show that wasting was not associated with the dietary diversity levels of the derived pattern at $p < 0.05$. The dietary diversity levels were fitted with demographic factors as presented in table 8.

Table 8: Logistic regression model for wasting with demographic factors

	Estimate	**Z value**	$Pr(>\|z\|)$
(Intercept)	-3.271	-6.562	$5.31e-11$
Dietary diversity level			
Low (ref)			
Medium	-0.402	-1.742	0.082
High	-0.085	-0.393	0.695
Sex of household head	0.070	0.739	0.460
Male (ref)			
Female	0.029	0.139	0.889
Region			
North (ref)			
Central	0.600	1.658	0.097
South	0.854	2.462	0.014
Place of residence			
Urban (ref)			
Rural	-0.245	-0.957	0.338
Highest Level of education			
No education (ref)			
Primary	0.038	0.134	0.893
Secondary	-0.123	-0.336	0.737
Tetiary	-0.687	-0.633	0.527
Wealth Index			
Poorest (ref)			
Poorer	-0.048	-0.195	0.846
Middle	-0.21442	-0.782	0.434
Richer	-0.388	-1.241	0.215
Richest	-0.518	-1.398	0.162

AIC: 1037.7

The study results in 8 show that region is associated with wasting. Children residing in the southern region were likely to be wasted at 5% significance level.

The wasting model with AIC 1032.1 fitted with dietary patterns only is a better model compared to the one fitted with demographic factors with AIC 1037.7.

4.2.2.3 Underweight binary logistic regression models for PCR dietary pattern

Table 9: Logistic regression model for underweight

	Estimate	**z value**	$Pr(>\|z\|)$
(Intercept)	-2.225	-20.929	$< 2e-16$
Dietary diversity score			
Low (ref)			
Medium	0.036	0.239	0.811
High	0.099	0.665	0.506

AIC: 1959.3

The binary logistic regression output for underweight in 9 shows that there is no linkage between underweight and the dietary diversity levels since no level was significant at P< 0.05.

Table 10: Logistic regression model for underweight with demographic factors

	Estimate	**Z value**	$\Pr(>\lvert z\rvert)$
(Intercept)	-2.609	-7.814	$5.55e-15$
Dietary diversity level			
Low (ref)			
Medium	0.012	0.081	0.936
High	0.182	1.206	0.228
Sex of household head			
Male (ref)			
Female	0.125	0.902	0.367
Region			
North (ref)			
Central	0.280	1.351	0.177
South	0.301	1.501	0.133
Place of residence			
Urban (ref)			
Rural	0.575	2.718	0.007
Highest Level of education			
No education (ref)			
Primary	-0.029	-0.16	0.873
Secondary	-0.276	-1.122	0.262
Tetiary	-1.108	-1.061	0.289
Wealth Index			
Poorest (ref)			
Poorer	-0.223	-1.397	0.163
Middle	-0.554	-2.979	0.002
Richer	-0.660	-3.194	0.001
Richest	-0.793	-3.251	0.001

AIC: 1933

The model was re-fitted with demographic factors as presented in table 10. The results show that place of residence and wealth index are associated with underweight at $p<0.05$. Children from the poor wealth profile were likely to be stunted than those in the richer wealth profile. Underweight model with AIC 1933 fitted with both dietary pattern levels and demographic factors is a better model compared to the one fitted with dietary pattern only with 1959.3

4.3 Application of RRR to examine the relationship between nutritional outcomes and dietary patterns in Malawi using reduced rank regression

4.3.1 Application of reduced rank regression and logistic regression to assess the linkage between dietary patterns and undernutritional outcomes

Reduced rank regression of type "pca" was performed with full rank in R software version 3.6.1 to obtain components. Loadings of the RRR were obtained to identify food groups to be included in the derived components. Loadings for principal components and their associated food groups are highlighted in table 11

Table 11: Loadings of the Reduced Rank Regression

Variable	PC 1	PC 2	PC 3
Carbohydrates			
Bread, noodles, other food made from grains	**-0.298**	0.142	-0.29
Other solid and semi-solid food	-0.374	0.0377	**-0.803**
Soup/Clear broth	-0.297	**0.374**	0.183
Potatoes, cassava and other tubers	**-0.173**	-0.159	0.13
Fortified baby food (cerelac)	-0.0961	**-0.144**	0.0776
Legumes and Nuts			
Beans, Peas, lentils, nuts	-0.245	**-0.595**	0.0931
Fruits			
Mangoes, pawpaw and other vitamin A fruits	**-0.342**	0.176	0.254
Any other fruit	**-0.265**	-0.0822	0.174
Vegetables			
Dark green leafy vegetables	**-0.399**	0.439	0.188
Animal food			
Fish, shell fish	**-0.203**	-0.0939	0.025
Eggs	-0.161	-0.175	0.119
Meat (beef, pork, lamb, chicken)	-0.154	**-0.172**	0.0564
Fats			
Oil, fats, butter and products made from them	0.173	**-0.233**	-0.0504
Other Foods and Drinks			
Gave child other liquid	**0.248**	-0.0624	0.187
Juice	-0.0969	-0.0483	**0.112**
Chocolate, sweets, candies and pastries	-0.204	**-0.275**	-0.0671
Soft drinks	-0.0625	-0.0462	**0.0651**

Food groups associated with reduced rank components are highlighted in table 11. All the 3 principle components were used in regression modeling . PC1 , PC2 and PC 3 variables were split into categories of 3 dietary diversity levels namely low, medium and high. The food groups were categorised based on principal componets with highest loading. Food groups associated with PC1, PC2 and PC3 are presented in table **??**

Table 12: Showing reduced rank regression derived components and their associated food item

Component	Food Groups	Dietary Pattern
PC 1	Bread, noodles, other food made from grains potatoes, cassava and other tubers Mangoes, pawpaw and other vitamin A fruits Other fruits Dark green leafy vegetables Fish, shell fish Other liquid	Carbohydrates-fruits-vegetables-Fish
PC2	Soup/Clear broth Fortified baby food Beans, Peas, lentils, nuts Meat (beef, pork, lamb, chicken) Oils, fats, butter and products made from them eggs Chocolate, sweets, candies and pastries	Legumes-animal protein-sugar
PC3	Other solid and semi-solid food juice soft drinks	Solid-foods-other-liquid

4.3.2 Logistic regression modelling for reduced rank regression derived dietary patterns

4.3.2.1 Stunting binary logistic regression models for RRR dietary patterns

Table 13: Logistic regression model for stunting

	Estimate	**z value**	$Pr(>\lvert z\rvert)$
(Intercept)	-1.02305	-8.314	$< 2e-16$
PC1			
Low (ref)			
Medium	0.19088	1.967	0.0492
High	-0.20356	-1.55	0.1212
PC2			
Low (ref)			
Medium	0.10575	0.881	0.3785
High	0.26034	2.243	0.0249
PC3			
Low (ref)			
Medium	-0.04168	-0.356	0.7221
High	-0.12907	-1.272	0.2035

AIC: 3499.3

Table 13 results show that the Carbohydrates-fruits-vegetables-Fish pattern was associated with stunting at 5% significance level with the medium diversity level having (OR=1.21, P< 0.05). Furthermore the legumes-animal protein-sugar dietary pattern was associated with stunting at 5% significance level with the high diversity level having (OR=1.21, P< 0.05). This implies that underfive children who were not consuming food groups in Carbohydrates-fruits-vegetables-Fish pattern and legumes-animal protein-sugar pattern were likely to be stunted.

Table 14: Logistic regression model for stunting with demographic factors

	Estimate	**Z value**	$\Pr(>\lvert z\rvert)$
(Intercept)	-1.035	-4.259	$2.05e-05$
PC1			
Low (ref)			
Medium	0.135	1.364	0.173
High	-0.288	-2.15	0.032
PC2			
Low (ref)			
Medium	0.057	0.463	0.643
High	0.180	1.532	0.126
PC3			
Low (ref)			
Medium	-0.030	-0.254	0.799
High	-0.096	-0.926	0.354
Sex of household head			
Male (ref)			
Female	0.056	0.586	0.558
Region			
North (ref)			
Central	0.305	2.29	0.022
South	0.247	1.917	0.055
Place of residence			
Urban (ref)			
Rural	0.308	2.349	0.019
Highest Level of education			
No education (ref)			
Primary	-0.209	-1.666	0.096
Secondary	-0.494	-3.017	0.003
Tetiary	-0.808	-1.713	0.087
Wealth Index			
Poorest (ref)			
Poorer	-0.052	-0.455	0.649
Middle	-0.233	-1.859	0.063
Richer	-0.424	-3.086	0.002
Richest	-0.444	-2.838	0.005

AIC: 3460.4

The model was fitted with demograpic factors and the results in table 14 reveals Carbohydrates-fruits-vegetables-Fish pattern was associated with stunting with the medium dietary diversity level $p< 0.05$. Further , the demographic factors region , type of residence, education and wealth index are associated with stunting at 5% significance level. Model comparison using AIC show that the model fitted with demographic factors is a better model with AIC 3460.4 compared to the model fitted with dietary patterns only with AIC 3499.3 .

4.3.2.2 Wasting binary logistic regression models for RRR dietary patterns

Table 15: Logistic regression model for wasting

	Estimate	**z value**	$Pr(>\lvert z\rvert)$
(Intercept)	-3.61475	-11.283	$< 2e-16$
PC1			
Low (ref)			
Medium	-0.08094	-0.339	0.7343
High	0.2894	1.033	0.3016
PC2			
Low (ref)			
Medium	0.59733	2.189	0.0286
High	0.25554	0.865	0.3872
PC3			
Low (ref)			
Medium	0.48195	1.883	0.0598
High	-0.09165	-0.359	0.7194

AIC : 1030.4

On Wasting, the second dietary pattern was associated with wasting given the medium dietary diversity level having $p< 0.05$ (OR=1.817) as presented in table 15. on the contrary, the first (Carbohydrates-fruits-vegetables-Fish pattern) and the third (Solid-foods-other-liquid pattern) were not associated with wasting. The wasting model fitted with demographic factors with

Table 16: Logistic regression model for wasting with control factors

	Estimate	**Z value**	Pr($>\|z\|$)
(Intercept)	-3.930	-6.757	$1.41e-11$
PC1			
Low (ref)			
Medium	-0.116	-0.479	0.632
High	0.207	0.729	0.466
PC2			
Low (ref)			
Medium	0.615	2.246	0.025
High	0.208	0.702	0.483
PC3			
Low (ref)			
Medium	0.615	2.246	0.025
High	0.208	0.702	0.483
Sex of household head			
Male (ref)			
Female	0.034	0.164	0.870
Region			
North (ref)			
Central	0.632	1.741	0.082
South	0.902	2.593	0.010
Place of residence			
Urban (ref)			
Rural	-0.274	-1.066	0.286
Highest Level of education			
No education (ref)			
Primary	0.032	0.112	0.911
Secondary	-0.109	-0.296	0.767
Tetiary	-0.741	-0.682	0.495
Wealth Index			
Poorest (ref)			
Poorer	-0.039	-0.16	0.873
Middle	-0.172	-0.627	0.531
Richer	-0.374	-1.194	0.233
Richest	-0.473	-1.278	0.201

AIC: 1035.7

On fitting wasting model with control variables, the output in table 16 the second (Legumes-animal protein-sugar patterns) and the third (Solid-foods-other-liquid pattern) were associated with wasting. Region was also associated with wasting at 5% significance level. Wasting model fitted with dietary diversity levels only with AIC 1030.4 is a better model compared to the model fitted with odemographic factors with AIC 1035.7 .

4.3.2.3 Underweight binary logistic regression models for RRR dietary patterns

Table 17: Logistic regression model for underweight

	Estimate	**z value**	$Pr(>\|z\|)$
(Intercept)	-2.41851	-12.54	$< 2e-16$
PC1			
Low (ref)			
Medium	0.08244	0.568	0.57035
High	0.20349	1.049	0.29433
PC2			
Low (ref)			
Medium	0.29719	1.633	0.10245
High	0.47895	2.658	0.00785
PC3			
Low (ref)			
Medium	-0.15735	-0.904	0.36574
High	-0.17717	-1.179	0.23826

AIC: 1956

The second dietary diversity pattern legumes-animal protein-sugar pattern was associated with underweight with the high diversity level (OR= 1.614 , p< 0.05). The second dietary pattern comprised of food groups in table 12

Table 18: **Logistic regression model for underweight with demographic factors**

	Estimate	**Z value**	Pr(> $\|z\|$)
(Intercept)	-2.683	-7.134	$9.77e-13$
PC1			
Low (ref)			
Medium	-0.005	-0.032	0.974
High	0.112	0.567	0.570
PC2			
Low (ref)			
Medium	0.225	1.215	0.224
High	0.393	2.168	0.030
PC3			
Low (ref)			
Medium	-0.151	-0.855	0.393
High	-0.138	-0.901	0.367
Sex of household head			
Male (ref)			
Female	0.118	0.855	0.393
Region			
North (ref)			
Central	0.280	1.344	0.179
South	0.302	1.501	0.133
Place of residence			
Urban (ref)			
Rural	0.559	2.638	0.008
Highest Level of education			
No education (ref)			
Primary	-0.037	-0.204	0.838
Secondary	-0.259	-1.05	0.294
Tetiary	-1.027	-0.982	0.326
Wealth Index			
Poorest (ref)			
Poorer	-0.217	-1.357	0.175
Middle	-0.539	-2.889	0.004
Richer	-0.637	-3.074	0.002
Richest	-0.745	-3.052	0.002

AIC:1934.9

Table 18 presents the output for underweight and demographic factors. The dietary diversity levels for the second pattern were found to be associated with underweight when fitted with demographic factors. More over place of residence and wealth index was associated with underweight at 5% significance level. Underweight model fitted with dietary diversity levels of the dietary pattern and demographic factors with AIC 1934.9 is a better model compared to the model with dietary diversity levels only with AIC 1956.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATION

5.1 Introduction

This chapter of the thesis focuses on the overall discussion inline with the study objectives, conclusion drawn from the discussion and recommendations made based on the conclusions. This study aimed at modelling the linkage between dietary patterns and undernutrition outcomes. Prior to modelling, dietary patterns were derived using principal component regression and reduced rank regression. The models were fitted using dietary diversity levels of the dietary patterns as explanatory variables for stunting, wasting and underweight. In addition, demographic factors associated with undernutrition were investigated.

5.2 Discussion

The study found that 57% of the under five children in Malawi consume only 1 to 3 food groups belonging to low dietary diversity and only 17% consume 6 to 12 food groups thus high dietary diversity level as shown in table 3. Furthermore children residing in rural areas have low consumption percentage for most of the food groups compared to those residing in urban areas due to limited access to food. This is also explained in the Malawi nutrition profile (2018) that people living in rural area have unreliable livelihood options that limit their access to food resulting into undernutrition.

Principal component regression derived dietary pattern was a carbohydrates-legumes-fruits-vegetables-animal-protein-fats-sugar dominated pattern. The pattern was split into dietary diversity levels. The results in table 5 , table 7 and table 9 show that dietary patterns are associated with stunting. This agrees with (Batis et al., 2016) and Machira and Chirwa (2020) that fruits and vegetables, carbohydrates dominated dietary patterns are associated with stunting. It is imperative to note that demographic factors including place of residence, parent or guardian education level, wealth index and region are also associated with stunting as presented in 6. This result is inline with (Ntenda & Chuang, 2018) who explains that social demographic factors effect nutritional status of people including underfive children.

Contrary to stunting, the derived dietary patterns were not associated with wasting and underweight. However when fitted with the demographic factors region of residence was associated with wasting. On the other hand, residence and wealth index were associated with underweight at 5% significance level. This shows that when implementing interventions to fight wasting and underweight in underfive children type of residence and economic factors have to be put into consideration as reiterated by the (UNICEF & WHO, 2008). Further, unbalanced dietary food consumption for underfive children increase vulnerability to various diseases and mortality experiences in worst case scenario (Nyanhanda et al., 2023)

Reduced rank regression model yielded three dietary patterns namely Carbohydrates-fruits-vegetables-Fish pattern, Legumes-animal protein-sugar pattern and Solid-foods-other-liquid pattern presented in table 12. From the results in table 13 at Carbohydrates-fruits-vegetables-Fish pattern and the Legumes-animal protein-sugar pattern were associated with stunting among under five children. This is also explained by (Meshram et al., 2012) that children who consume less of the food groups in the mentioned dietary patterns are associated with stunting. Lin et al. (2007) also adds on describing carbohydrates, fruits, vegetables and and proteins as important to be considered in consumption when addressing stunting. UNICEF and WHO (2008) explains that combining dietary food groups is important in combating undernutrition to reach a recommended consumption of food groups for child nutritional health. Fitting the stunting model with demographic factors table 14 shows that the study found that stunting was associated with education level of the parent or guardian, region education and wealth index. This concurs with Mphamba et al. (2024) who explains that education and wealth are among the main contributors to inequality of opportunity in under-five nutrition causing children from parents with low education and belonging to low wealth index being stunted. This was also found in a study done by (Ntenda & Chuang, 2018) where the study found that demographic factors have an impact in the nutritional health status of underfive children hence these factors have to be incoporated for successful implementation of undernutrition interventions.

The study results also indicate that the two dietary patterns derived from RRR were associated with wasting as presented in table 15. Children who were not consuming carbohydrates, fruits fish and liquids had higher probability of being wasted. Yue et al. (2022) explains that nutritional deficiencies which result from eating unbalanced diet exacerbates undernutrition in underfive children. In, addition region of residence for the under five children was associated with wasting. The Malawi nutritional profile report (2018) indicates that children in the southern region of Malawi are assoicated with wasting compared to those in the northern region.

Underweight is also one of the most prominent type of undernutrition (Mank et al., n.d.). From the results in table 17 underweight was associated with the legumes-animal protein-sugar pattern.Further Ntenda (2019) recommended that, underweight can be treated by implementing interventions that aim at improving the growth and development of children in their early years including dietary diversification. Moreover, table 18 place of residence and wealth index was also associated with underweight. Children residing in rural areas in Malawi were associated with being underweight. According to Machira and Chirwa (2020) underfive children residing in rural areas are associated with underweight because of limited access to diversified food.

5.3 Conclusion

The current study has found that there is a linkage between dietary patterns derived from PCR and RRR and undernutrition outcomes. The dietary patterns derived from the two methods include almost similar food groups associated with stunting, wasting and underweight. The similar food groups include carbohydrates, legumes, fruits, vegetables, animal protein, fats, and sugar hence using both principal component regression and reduced rank regression is important in assessing dietary patterns and associated outcomes. Further dietary diversity has an effect in undernutrition, medium and high dietary diversity levels were significant in explaining stunting, wasting and underweight at 5% significance level. This implies that children whose dietary consumption was not diversified were likely to be undernourised.

Furthermore, the study also found that demographic factors such as parents or guardian education level, residence , region and economic status are also associated with undernutrition. Moreover, in the current study, models fitted with both dietary patterns and demographic factors were found to be better models using AIC model selection compared to those fitted with dietary patterns only. However, the study results show that sex of the household head was not associated with stunting , underweight and wasting in all the models fitted with demographic factors for both PCR and RRR. Therefore there is need to enhance and promote dietary diversification among underfive children especially among rural and low income households. In addition, public health education would enhance knowledge on importance of dietary practices among rural women who mostly take care of the underfive children. The study results will help nutrition and health agencies to reach out to underfive children in Malawi with more specific undernutrition interventions and reduce stunting, wasting and underweight levels hence contributing to the economic growth of Malawi as a country.

5.4 Recommendation

The current study findings support that dietary diversification improvement should be one of the measures to improve stunting, wasting and underweight in under-five children. Futher, demographic factors plays a low in the nutritional status of underfive children. Therefore reducing undernutrition requires a multi-sectoral collaborative efforts such as promoting nutrition-sensitive agricultural interventions and promoting balanced diet through health education using community-based approaches, and incorporating demographic factors such as education level, wealth index and place of residence can help improve nutritional status of underfive children in Malawi.

REFERENCES

Ali, A., Margetts, B. M., & Zainuddin, A. A. (2021). *Exploration of the principal component analysis (pca) approach in synthesizing the diet quality of the malaysian population* (Vol. 13). doi: 10.3390/nu13010070

Association), W. N. (2021). *The un decade of nutrition: Wphna's position at mid-term.*

Bank, W. (2016). *World bank helps malawi control malnutrition in under-five children.* Retrieved from `www.worldbank.org/en/news/press-release/2016/06/23/world-bank-helps-malawi-control-malnutrition-in-under-five-children`

Batis, C., Mendez, M. A., Gordon-Larsen, P., Sotres-Alvarez, D., Adair, L., & Popkin, B. (2016). Using both principal component analysis and reduced rank regression to study dietary patterns and diabetes in chinese adults. *Public Health Nutrition*, *19*, 195-203. doi: 10.1017/S1368980014003103

Bilinsky, P., & Swindale, A. (2006). Household dietary diversity score (hdds) for measurement of household food access : Indicator guide version 2 anne swindale household dietary diversity score (hdds) for measurement of household food access : Indicator guide version 2.

Braeken, J., & Assen, M. A. V. (2017). An empirical kaiser criterion. *Psychological Methods*, *22*, 450-466. doi: 10.1037/met0000074

Centre, U. N. (2015). *Un adopts new global goals, charting sustainable development for people and planet by 2030.* Retrieved from `http://www.un.org/en/development/desa/news/sustainable/un-adopts-new-global-goals.html#more-15178`

Curtis, A. R., Livingstone, K. M., Daly, R. M., Brayner, B., Abbott, G., & Kiss, N. (2023). *Dietary patterns, malnutrition, muscle loss and sarcopenia in cancer survivors: findings from the uk biobank.* doi: 10.1007/s11764-023-01428-8

Fariq, R., & Fatah, F. K. (2016). (pdf) parameter estimation for binary logistic regression using different iterative methods.

Grewall, M., & Wakim, S. (2021). $4.6_undernutrition - biologylibretexts$.

Guide, I. (2012). Interpretation guide. *Nutrition Landacape Information System (NLIS)*, 1-51. doi: 10.1159/000362780.Interpretation

Harris, K. J. (2021). Primer on binary logistic regression - pmc.

HLPE. (2017). High level panel of experts. 2017. nutrition and food systems. *Committee o World Food Security (CFS)*, *44*, 150. Retrieved from http://www.fao.org/3/a-i7846e.pdf

Huybrechts, I., Lioret, S., Mouratidou, T., Gunter, M. J., Manios, Y., Kersting, M., . . . McNaughton, S. A. (2017). *Using reduced rank regression methods to identify dietary patterns associated with obesity: A cross-country study among european and australian adolescents* (Vol. 117). doi: 10.1017/S0007114516004669

IMF. (2007). *Malawi: Poverty reduction strategy paper-growth and development strategy.*

Kassie, G. W., & Workie, D. L. (2020). *Determinants of under-nutrition among children under five years of age in ethiopia* (Vol. 20). doi: 10.1186/s12889-020-08539-2

Kelly, P. (2011). *Undernutrition* (Vol. 18). doi: 10.1002/9781444327779.ch17

Khamis, A. G., Mwanri, A. W., Ntwenya, J. E., & Kreppel, K. (2019). The influence of dietary diversity on the nutritional status of children between 6 and 23 months of age in tanzania. *BMC Pediatrics*, *19*, 1-9. doi: 10.1186/s12887-019-1897-5

Kramer, R. (1998). *Principal component regression.* doi: 10.1201/9780203909805.ch7

Krebs-Smith, S. M., & Hoffman, K. (2014). Approaches to dietary pattern analyses : Potential to inform guidance. *2015 Dietary Guidelines Advisory Council Second Meeting*, 1-67. Retrieved from https://health.gov/dietaryguidelines/2015-binder/meeting2/day1Agenda.aspx%0Ahttps://health.gov/dietaryguidelines/2015-binder/meeting2/docs/workGroupPresentations/DGAC_Patttern_Sue_Krebs_Smith_2-27-14.pdf

Lin, C. A., Boslaugh, S., Ciliberto, H. M., Maleta, K., Ashorn, P., Briend, A., & Manary, M. J. (2007). *A prospective assessment of food and nutrient intake in a population of malawian children at risk for kwashiorkor* (Vol. 44). doi: 10.1097/

MPG.0b013e31802c6e57

Machira, K., & Chirwa, T. (2020). *Dietary consumption and its effect on nutrition outcome among under-five children in rural malawi* (Vol. 15). doi: 10.1371/journal.pone.0237139

Malawi-NSO, & ICF. (2017). *Malawi demographic and health survey 2015-16.*

Maleta, K. (2006). *Undernutrition - pmc.*

Mank, I., Belesova, K., JanIssouf, I. B., Traoré, Wilkinson, P., Danquah, I., & Sauerborn, R. (n.d.). The impact of rainfall variability on diets and undernutrition of young children in rural burkina faso - pubmed. *2021*.

Mayasari, N. R., Bai, C. H., Chao, J. C. J., Chen, Y. C., Huang, Y. L., Wang, F. F., ... Chang, J. S. (2023). *Relationships between dietary patterns and erythropoiesis-associated micronutrient deficiencies (iron, folate, and vitamin b12) among pregnant women in taiwan* (Vol. 15). doi: 10.3390/nu15102311

Meshram, I. I., Arlappa, N., Balakrishna, N., Rao, K. M., Laxmaiah, A., & Brahmam, G. N. V. (2012). *Trends in the prevalence of undernutrition, nutrient food intake and predictors of undernutrition among under five year tribal children in india* (Vol. 21).

Moradi, M., Jalilpiran, Y., Askari, M., Surkan, P. J., & Azadbakht, L. (2022). *Associations between mother–child dyad dietary patterns and child anthropometric measures among 6-year-old children* (Vol. 181). doi: 10.1007/s00431-021-04180-2

Movement, S. (2015). *Malawi - sun.*

Mphamba, P., Chirwa, G., & Mazalale, J. (2024). An evolution of inequality of opportunity in the nutritional outcomes of under-five children in malawi - pmc.

Mukherjee, A. (2013). Topics on reduced rank methods for multivariate regression. Retrieved from `http://deepblue.lib.umich.edu/handle/2027.42/99837`

Ntenda, P. A. M. (2019). *Association of low birth weight with undernutrition in* *preschool-aged children in malawi* (Vol. 18). doi: 10.1186/s12937-019-0477-8

Ntenda, P. A. M., & Chuang, Y. C. (2018). *Analysis of individual-level and community-level effects on childhood undernutrition in malawi* (Vol. 59). doi: 10.1016/j.pedneo.2017.11.019

Nunget, R., Levin, C., Harry, J., & Hutchinson, B. (2019). Economic effects of the double burden of malnutrition - the lancet.

Nyanhanda, T., Mwanri, L., & Mude, W. (2023, 5). Double burden of malnutrition$_{a}$*populationlevelcom*

sectionalstudyacrossthreesub−saha.Environ Res Public Health.Retrievedfrom

of Nutrition HIV, D., & AIDS. (2018). *National multi-sector nutrition policy 2018-2022*.

Parry, S. (2020). *Cornell statistical consulting unit ordinal logistic regression models and statistical software: What you need to know statnews 91.*

Profile, . M. N. (2018). Zimbabwe : Nutrition profile. , 1-6.

Report, G. N. (2020). *Global nutrition report.*

Report, G. N. (2021). Global nutrition report: the state of global nutrition. executive summary. *Global Nutrition Report*. Retrieved from https://globalnutritionreport.org/reports/2021-global-nutrition-report/

Report, T. G. N. (2018). *Global nutrition report: Shining a light to spur action on nutrition.* Retrieved from http://www.segeplan.gob.gt/2.0/index.php?option=com_content&view=article&id=472&Itemid=472

Sambo, T. A., Oguttu, J. W., & Mbombo-Dweba, T. P. (2022). *Analysis of the dietary diversity status of agricultural households in the nkomazi local municipality, south africa* (Vol. 11). doi: 10.1186/s40066-022-00387-0

Sanchez, G., & Marzban, E. (2020). *15 principal components regression* $_{a}$*llmodelsarewrong*$_{c}$*once*

Schwerin, H. S., Stanton, J. L., Riley, A. M., Schaefer, A. E., Leveille, G. A., Elliott, J. G., . . . Brett, B. E. (1981). Food eating patterns and health: a reexamination of the ten-state and hanes i surveys *(Vol. 34). doi: 10.1093/ajcn/34.4.568*

Shrestha, A., Shrestha, A., & Cissé, G. (2021). Dietary patterns measured by principal component analysis and its association with stunting among nepalese schoolchildren in nepal *(Vol. 19). doi: 10.3126/kumj.v19i1.49528*

Silvera, S. A. N., Mayne, S. T., Risch, H. A., Gammon, M. D., Vaughan, T., Chow, W. H., . . . Blot, W. J. (2011). Principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer *(Vol. 21). doi: 10.1016/j.annepidem.2010.11.019*

Social, T. (2020). The cost of in malawi in malawi.

Soetewey, A. (2024). Binary logistic regression.

Tesfay, F., Javanparast, S., Gesesew, H., Mwanri, L., & Ziersch, A. (2022). Original research$_{c}$*haracteristicsandimpactsofnutritionalprogrammestoaddressundernutritionofa saharanafrica*$_{a}$*systematicreviewofevide.*

UNICEF, & WHO. (2008). Indicators for assessing infant and young child feeding practices. World Health Organization, WHA55 A55/*, 19. Retrieved from* `http://apps.who.int/iris/bitstream/handle/10665/44306/9789241599290_eng.pdf?sequence=1%0Ahttp://whqlibdoc.who.int/publications/2008/9789241596664_eng.pdf%5Cnhttp://www.unicef.org/programme/breastfeeding/innocenti.htm%5Cnhttp://innocenti15.net/declaration.`

(WHO), W. H. O. (2020). Fact sheets: Malnutrition. *Retrieved from* `https://www.who.int/news-room/fact-sheets/detail/malnutrition`

Wiesmann, D., Bassett, L., Benson, T., & Hoddinott, J. (2009). Validation of the world food programme's food consumption score and alternative indicators of household food security. IFPRI Discussion Paper, 00870, *1-105.*

Winter. (2024). Newsom psy 522/622 multiple regression and multivariate quantitative methods, winter 2024 1.

Yue, T., Zhang, Q., Li, G., & Qin, H. (2022). Global burden of nutritional deficiencies among children under 5 years of age from 2010 to 2019 - pmc.

Zhen, S., Ma, Y., Zhao, Z., Yang, X., & Wen, D. (2018). Dietary pattern is associated with obesity in chinese children and adolescents: Data from china health and nutrition survey (chns) *(Vol. 17). doi: 10.1186/s12937-018-0372-8*

Zheng, B., Liu, Q., Lyu, J., & Yu, C. (2022). Introduction of reduced rank regression and development of a user-written stata package *(Vol. 43). doi: 10.3760/cma.j.cn112338-20210222-00136*

APPENDICES

Appendix A: R code used for Principal Component Regression, Reduced Rank Regression and Binary Logistic Regression Modelling

```
# Importing cleaned data set from Stata


install.packages("rlang")

library("rlang")

library(haven)


DataMayFinalCleaned <- read_dta("F:/MSc School

Books & Modules/Thesis/Thesis Write-up/January

2022 Revised Thesis Copy - Nov 9,

2022 - April - May 2023/DataMayFinalCleaned.dta")

View(DataMayFinalCleaned)


library("tidyverse")

library(tidyverse)


# Creating a Subset data of explanatory

variables (foods/liquids)

attach(DataMayFinalCleaned)

MyMayExplanatoryData<-DataMayFinalCleaned[,

c("V414E", "V414F", "V414S", "V412C",

"V412A", "V414O", "V414K", "V414L",
```

```
"V414J", "V414G", "V414H", "V414N", "V414Q",
 "V414R","V410", "V413", "V413A")]
library(tidyr)
library("dplyr")

MyMayExplanatoryData<- rename
(MyMayExplanatoryData, Bread_Noodles="V414E", Potatoes_Cassava="V414F",
Solid_food="V414S",Soup="V412C", Cerelac="V412A",
Beans="V414O", Mangoes_Pawpaws="V414K",
Other_Fruits="V414L", Leaf_Vegetables="V414J",
Eggs="V414G", Meat="V414H", Fish="V414N",
 oil="V414Q", Chocolate="V414R",Juice="V410",
 Other_liquid="V413", Soft_Drinks="V413A")

view (MyMayExplanatoryData) # for
explanation of the principle components



# Creating a response variable
ResponseData<-DataMayFinalCleaned[,c("CASEID",
 "B4", "HW1", "HW2", "HW3")] # Creating subset
  data of response variables
ResponseData<- rename(ResponseData,
CaseID="CASEID", SexC="B4",
```

```
 Age="HW1",Weight="HW2", Height="HW3")

view(ResponseData)



# coding missing values in the response
data set(Age , Weight and Height variables)
 for respondent not present 9994, respondent
 refused 9995 and Other 9996
ResponseData$Age[ResponseData$Age==9994] <- NA
ResponseData$Age[ResponseData$Age==9995] <- NA
ResponseData$Age[ResponseData$Age==9996] <- NA

ResponseData$Weight[ResponseData$Weight==9994]<- NA
ResponseData$Weight[ResponseData$Weight==9995]<- NA
ResponseData$Weight[ResponseData$Weight==9996]<- NA

ResponseData$Height[HDDSResponseData$Height==9994]<- NA
ResponseData$Height[HDDSResponseData$Height==9995]<- NA
ResponseData$Height[HDDSResponseData$Height==9996]<- NA



# calculating stunting, wasting and underweight using WHO standards
library(dplyr)
```

```
library(zscorer)


ResponseData$Age <- ResponseData$Age *(365.25/12)
#- For changing age into days
ResponseData$Height <- ResponseData$Height*(1/10)
#For changing child's height from centimeters to metres
ResponseData$Weight <- ResponseData$Weight*(1/10)
 #For changing child's Weight from current unit to kilograms


ResponseData<-addWGSR(data = ResponseData, sex = "SexC",
firstPart = "Height", secondPart = "Age", index = "hfa")
ResponseData<-addWGSR(data = ResponseData, sex = "SexC" ,
firstPart = "Weight", secondPart = "Age", index = "wfa")
ResponseData<-addWGSR(data = ResponseData, sex = "SexC",
firstPart = "Weight", secondPart = "Height", index = "wfh")


# Categorising Stunting , Wasting and
Underweight based on WHO definitions
ResponseData<-mutate(ResponseData,
 Stunting=ifelse(hfaz<(-2), 1, 0),
Wasting=ifelse(wfhz<(-2), 1, 0)
)


view(ResponseData)
```

```
#Making response variables to be recognized as categorical variables

ResponseData$Stunting <- as.factor(ResponseData$Stunting)
ResponseData$Wasting <- as.factor(ResponseData$Wasting)
ResponseData$Underweight <- as.factor(ResponseData$Underweight)



# Creating dataset for control variables
ControlData<-DataMayFinalCleaned[,c( "V151", "V101", "V102","V106", "V190A")]

ControlData<- rename(ControlData,
Sex_of_Household_Head="V151", Region="V101",Place_of_Residence="V102",
Highest_Education_Level="V106",Wealth_Index= "V190A")

view(ControlData)



#Making control variables to be recognized as categorical variables

ControlData$Sex_of_Household_Head <- as.factor(ControlData$Sex_of_Household_Head
ControlData$Region <- as.factor(ControlData$Region)
ControlData$Place_of_Residence <- as.factor(ControlData$Place_of_Residence)
ControlData$Highest_Education_Level <- as.factor(ControlData$Highest_Education_L
```

```
ControlData$Wealth_Index <- as.factor(ControlData$Wealth_Index)


#******Principle Component Regression ********

PCAFoods<-princomp(MyMayExplanatoryData,
scores = TRUE, cor = TRUE)
summary(PCAFoods)

MyMayExplanatoryData.PCAFoods <-
prcomp(MyMayExplanatoryData,
center=TRUE, scale.=TRUE)
summary(MyMayExplanatoryData.PCAFoods)
PCAFoods_Data <- as.data.frame
(MyMayExplanatoryData.PCAFoods$x)
PCA1<- PCAFoods_Data[, 1:3]
view (PCA1)

#Loadings of Principal Components
 and Sreeplot of eigen values
PCAloading <-loadings(PCAFoods)

#Scree Plot of Eigen Values
screeplot(PCAFoods, type="line", main = "scree plot")
```

```
#Split variable PC1 into 3 dietary levels


library(dplyr)


PCA1$group <- cut_number(PCA1$PC1, 3)

PCA1$group <- as.numeric(cut_number(PCA1$PC1, 3))


#Making PC1 variable to be recognized

as categorical a variable


PCA1$group<- as.factor(PCA1$group )

PCA1Table2<-  table(PCA1$group)




#Combining explanatory data set

and response data set and control

 Variables data set to be one data set

PCRModelData <- cbind(PCA1, ResponseData, ControlData)



# Model fitting with stunting, wasting and underweight

# Logistic Regression
```

```
# Stunting logistic regression model

PCRStuntingModel <- glm(Stunting ~ group ,

data = PCRModelData, family = 'binomial'

( link ='logit'))

summary(PCRStuntingModel)


exp(coef(PCRStuntingModel))[coef(summary(PCRStuntingModel))

[, "Pr(>|z|)"] < 0.05]

#for printing odds for only p-value <0.05



#  Stunting logistic regression model

with control variables

PCRStuntingModelControl <- glm(Stunting ~ group +

Sex_of_Household_Head + Region

      + Place_of_Residence + Highest_Education_Level

       + Wealth_Index , data = PCRModelData,

       family = 'binomial'( link ='logit'))


summary(PCRStuntingModelControl)



# Wasting logistic regression model
```

```
PCRWastingModel <- glm(Wasting ~ group ,
data = PCRModelData, family = 'binomial'
( link ='logit'))
summary(PCRWastingModel)
exp(coef(PCRWastingModel))[coef(summary(PCRWastingModel))
[, "Pr(>|z|)"] < 0.05]
#for printing odds for only p-value <0.05



# Wasting logistic regression model
with control variables
PCRWastingModelControl <- glm(Wasting ~ group +
Sex_of_Household_Head
+ Region
 + Place_of_Residence + Highest_Education_Level
  + Wealth_Index , data = PCRModelData,
   family = 'binomial'( link ='logit'))

summary(PCRWastingModelControl )



# Underweight logistic regression model
PCRUnderweightModel <- glm(Underweight  ~ group,
 data = PCRModelData, family = 'binomial'( link ='logit'))
```

```
summary(PCRUnderweightModel)

exp(coef(PCRUnderweightModel))[coef(summary(PCRUnderweightModel))
[, "Pr(>|z|)"] < 0.05] #for printing odds for
only p-value <0.05


# Underweight logistic regression model
 with control variables
PCRUnderweightModelControl <- glm(Underweight ~
 group + Sex_of_Household_Head + Region
   + Place_of_Residence + Highest_Education_Level
    + Wealth_Index , data = PCRModelData,
    family = 'binomial'( link ='logit'))

summary(PCRUnderweightModelControl)


# *********************Reduced Rank Regression*********************
library(rrr)

view (ResponseData)
view (RRRExplanatoryData)
```

```
x <- MyMayExplanatoryData[,1:17]

y <-RRRResponseData[,c( "Stunting", "Wasting", "Underweight")]



RRRResults<-rrr(x, y, type = "pca", rank = "full", k = 0)

RRRPC_Data<-scores(x, y, type = "pca", rank = "full", k = 0)



#Split variables PC1 , PC2 and PC3 into groups


#Splitting PC1


RRRPC_Data$PC1Group <- cut_number(RRRPC_Data$PC1, 3)

summary(RRRPC_Data$PC1Group)

RRRPC_Data$PC1Group <- as.numeric(cut_number(RRRPC_Data$PC1, 3))


#Making PC1 variable to be recognized as categorical a variable


RRRPC_Data$PC1Group<- as.factor(RRRPC_Data$PC1Group)


summary(RRRPC_Data$PC1Group)
```

```
#Splitting PC2

RRRPC_Data$PC2Group <- cut_number(RRRPC_Data$PC2, 3)
summary(RRRPC_Data$PC2Group)
RRRPC_Data$PC2Group <- as.numeric(cut_number(RRRPC_Data$PC2, 3))

#Making PC2 variable to be recognized as categorical a variable
RRRPC_Data$PC2Group<- as.factor(RRRPC_Data$PC2Group)

summary(RRRPC_Data$PC2Group)


#Splitting PC3

RRRPC_Data$PC3Group <- cut_number(RRRPC_Data$PC3, 3)
summary(RRRPC_Data$PC3Group)
RRRPC_Data$PC3Group <- as.numeric(cut_number(RRRPC_Data$PC3, 3))

#Making PC3 variable to be recognized as categorical a variable
RRRPC_Data$PC3Group<- as.factor(RRRPC_Data$PC3Group)

summary(RRRPC_Data$PC3Group)
```

```
#Combining explanatory data set and
response data set and control Variables
data set to be one data set
RRRModelData <- cbind(RRRPC_Data, y, ControlData)




#Fitting logistic regression models for RRR
components results with PC1, PC2 , PC3 as
categorical variables with levels high, medium, low

#Fitting logistic regression models for RRR components
#Fitting RRR logistic model with stunting
RRRModelStunting <- glm(Stunting ~  PC1Group +
PC2Group + PC3Group,  data = RRRModelData,
family = 'binomial'( link ='logit'))
summary(RRRModelStunting )

exp(coef(RRRModelStunting))[coef(summary(RRRModelStunting))
[, "Pr(>|z|)"] < 0.05]
#for printing odds for only p-value <0.05
```

```
round(exp(cbind(OR = coef(RRRModelStunting), confint(RRRModelStunting))), 3)
#Odds ratio and confidence intervals




#Fitting RRR model  stunting with control variables
RRRStuntingModelControl <- glm(Stunting ~  PC1Group
+ PC2Group + PC3Group + Sex_of_Household_Head
 + Region + Place_of_Residence +
 Highest_Education_Level
 + Wealth_Index, data = RRRModelData,
 family = 'binomial'( link ='logit'))

summary(RRRStuntingModelControl)




#Fitting RRR model with wasting
RRRWastingModel <- glm(Wasting ~  PC1Group + PC2Group + PC3Group,
data = RRRModelData, family = 'binomial'( link ='logit'))
summary(RRRWastingModel)

exp(coef(RRRWastingModel))[coef(summary(RRRWastingModel))[,
```

```
"Pr(>|z|)"] < 0.05] #for printing odds for only p-value <0.05



#Fitting RRR model  wasting with control variables

RRRWastingModelControl <- glm(Wasting ~  PC1Group +

PC2Group + PC3Group + Sex_of_Household_Head

  + Region + Place_of_Residence + Highest_Education_Level

 + Wealth_Index, data = RRRModelData, family = 'binomial'( link ='logit'))

summary(RRRWastingModelControl)



#Fitting RRR model with underweight

RRRUnderweightModel <- glm(Underweight~  PC1Group + PC2Group

+ PC3Group, data = RRRModelData,

family = 'binomial'( link ='logit'))

summary(RRRUnderweightModel)

exp(coef(RRRUnderweightModel))

[coef(summary(RRRUnderweightModel))[, "Pr(>|z|)"] < 0.05]

#for printing odds for only p-value <0.05
```

```
#Fitting RRR model  wasting with control variables

RRRUnderweightModelControl <- glm(Underweight ~  PC1Group

+ PC2Group + PC3Group+

Sex_of_Household_Head

 + Region + Place_of_Residence + Highest_Education_Level

 + Wealth_Index, data = RRRModelData,

 family = 'binomial'( link ='logit'))

summary(RRRUnderweightModelControl)


# **************End of script**********************
```