Assessing the effectiveness of Lay Health Workers intervention on the management of hypertension in Agincourt rural, South Africa

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Student: **Munyaradzi Nyamukokoko (1816789)**

Supervisor 1: **Prof Eustasius Musenge**

Supervisor 2: **Prof Tobias Chirwa**

A Research Report Submitted to the Faculty of Health Sciences, University of the Witwatersrand in partial fulfilment of the requirements for the Degree of Masters in Epidemiology in the field of Biostatistics.

June 2022

Declaration

I, Munyaradzi Nyamukokoko, student No: **1816789** declare that this research report is my own work. It is being submitted in partial fulfilment of the degree of Master of Science in Epidemiology in the field of Biostatistics at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

Munyaradzi Nyamukokoko

March 2022

Dedication

Dedicating this dissertation to my mother and relatives who supported me all the way

Abstract

Background

Uncontrolled hypertension determined by systolic blood pressure greater than 140 mmHg and diastolic blood pressure greater than 90 mmHg is among the leading cause of death in the whole world. In Agincourt a sub district of Mpumalanga in South Africa, uncontrolled hypertension prevalence is very high approximately 40%. This figure is believed to be higher since almost half of the population does not know their blood pressure status. This might be due to poor management of chronic illness in this area. In this study, lay health workers (LHWs) were used in managing uncontrolled hypertension in Agincourt. The aim of this study was to assess the impact of using multilevel logistic regression models with random effects instead of the mixed effects model which was used in the primary study to assess the effectiveness of LHW intervention in the management of uncontrolled hypertension in Agincourt, South Africa.

Methods

The current study was a secondary data analysis of a cluster randomized controlled trial for data collected in Agincourt from 2013 to 2015. It adopted the primary study design as well as the sample size. The data consisted of population of hypertension patients aged 18 years and above. About 4000 people were allocated in each study arm which means that at least 500 people were allocated in each cluster. Demographic characteristics of the participants were shown as frequencies and percentages. The adjusted chi-squared test was used and the p-values were used to identify the association between uncontrolled hypertension and other factors. The treatment effect of lay workers was estimated using propensity score matching. Patient factors associated with the effectiveness of LHW in Agincourt were identified using level 1 (with individual level random effects) and level 2 (with both individual and cluster random effects) multilevel models.

Results

The descriptives statistics were focusing on the participants who were recorded both at baseline and at end of the intervention. The number of participants who were recorded both at baseline and end of intervention were 810 out of 4000. In the control arm, about 250 participants were recorded at baseline and 213 at the end of the intervention. In the intervention arm, 185 participants were recorded at baseline and 162 at the end of the intervention. There was no much difference on the social demographic variables between the control and the intervention arms at baseline and at the end of the intervention. Multilevel logistic regression model with individual random effects(level 1 multilevel model) was the best model compared to the model with both individual and cluster random effects (level 2 multilevel model) using intraclass correlation coefficient. The model highlighted that age, obesity, gender and smoking history were highly associated with uncontrolled hypertension. The odds (OR = 14.75; 95% CI: 8.1669 - 26.6513) of the elderly (40 years and above) of having uncontrolled hypertension is approximately seven times than that of young aged (18 - 39 years). The odds (OR = 2.8629; 95% CI: 1.5412 - 5.3178) of having uncontrolled hypertension among those who are obese is approximately three times than that of those who are not. The odds (OR = 2.2616; 95% CI: 0.9758 - 5.2414) of males of having uncontrolled hypertension is approximately twice than that of females.

Conclusion

The findings in this study reported that lay health workers intervention was not effective in reducing uncontrolled hypertension. The results align with those obtained in the primary study by Goudge et al. (2018). However, findings were made that LHW are important in undertaking some of the clinical duties such as measuring of blood pressure, booking of patient's appointments, retrieval of patient's files a day before the appointment and refilling afterwards, providing of health education and assisting nurses with pre-packing of medication (11).

Acknowledgements

I would like to thank my God for allowing me to undertake this research. Moreover, l would like to give thanks to my supervisors Prof Eustasius Musenge as well as Prof Tobias Chirwa. Without their guidance this research would not have been a success. I would like to express my gratitude and appreciation to Sub-Saharan African Consortium for Advanced Biostatistics (SSACAB) training programme for funding my studies. It has been an honour being considered part of the fellows in the SSACAB program. I also would like to thank all my friends for their support during the time I spent at the University of the Witwatersrand. Most importantly, l would like to give thanks to my family, none of this could have happened without them.

Contents

List of Figures

List of Tables

Acronyms

Definitions of terms

- **Controlled hypertension** - Systolic blood pressure which is less than 140 mmHg and diastolic blood pressure which is less than 90 mmHg.
- **Uncontrolled hypertension** - Systolic blood pressure greater than 140 mmHg and diastolic blood pressure greater than 90 mmHg.

Abbreviations

- SBP - Systolic blood pressure
- DBP - Diastolic blood pressure
- LHWs - Lay health workers
- SA - South Africa
- SSA - Sub Saharan Africa
- BP - Blood pressure
- CI - Confidence interval
- ICC - Intraclass correlation coefficient

Chapter 1

Introduction and Background

1.1 Outline of the Introduction chapter

In this introduction chapter, the rationale for this study was explained and an overview of the burden of hypertension globally and on a local scale in Africa was outlined. This research starts off by presenting the background of the study on which more emphasis is given on hypertension particularly in Agincourt, South Africa. This is then followed by problem statement that clearly outlines the problem at hand. It then proceeds to briefly describe the rationale or justification of the study, study aim and research questions. Finally, the chapter ends by describing the study objectives.

1.2 Background of the Study

Hypertension determined by systolic blood pressure (SBP) greater than 140 mmHg and diastolic blood pressure (DBP) greater than 90 mmHg is among the leading cause of death and disabilities in the whole world (1). Addo et al. (2007), pointed that hypertension has caused about 7.6 million deaths and contributed to 92 million disability adjusted life years (DALYs) worldwide in 2001. In another study, Limbani et al. (2019), highlighted that approximately 26.4% of the people from the age of 18 and above had hypertension, globally. The actual figure approximated is around 972 million people of the world's population with hypertension counting from the year 2000 (1). The projection of this global burden was expected to rise from 972 million to 1.6 billion by 2025 by Egan et al. (2019). Sub Saharan Africa due to her poverty is heavily affected by hypertension.

In Sub Saharan Africa (SSA), hypertension might be the leading cause of death amongst all other chronic illnesses across the continent. A study by Perkovic et al. (2007), hypertension has been considered a major public health problem in SSA. In a cross-sectional study by Gawattude et al. (2015) pointed out that out of the 650 million people in SSA, between 10 to 20 million may have hypertension and most of these people are not aware of their hypertension status. This might be so, because most of the SSA states are resource limited and most people don't have access to health services. Moreover, delivery of care is usually performed by lower-skilled healthcare workers, especially in rural areas, and access to medical care may require lots of money. South Africa is among the top countries that are heavily affected by hypertension worldwide.

According to the WHO-SAGE (World Health Organization Study on Global Ageing and Adult Health), South Africa has the highest prevalence of uncontrolled hypertension in Southern Africa. According to WHO data, approximately 27.4% of men and 26.1% of women in South Africa have hypertension, though prevalence of up to 60% has been reported (10). In South Africa, one of the districts with the highest prevalence of hypertension is Mpumalanga. Uncontrolled hypertension prevalence in Mpumalanga is slightly above 57% (11). Furthermore, Agincourt a subdistrict of Mpumalanga has prevalence of uncontrolled hypertension of approximately 30% (11). The estimate is believed to be higher since almost half of the population in Agincourt does not know their blood pressure. Goudge et al. (2018) argued that this might be due to poor management of chronic illnesses in this area. The condition only managed is less than 10% of the people with hypertension (11).

The general focus in chronic condition management is to improve patient care efficiency and ensure that the chronic conditions are well managed. However, the current nursing staff are already overburdened and struggle to effectively carry out these and provide proper consultations to the chronic patients (11). In the absence of an effective intervention strategy to control this epidemic, it is expected that uncontrolled hypertension prevalence will continue to rise leading to increase in mortality and disability. Therefore, the aim of this study was to assess the effectiveness of the Lay Health Worker intervention in the management of hypertension using multilevel logistic regression models with random effects.

1.3 Problem Statement

Researchers have discovered that mostly in Africa, there is a serious shortage of health workers (3). In South Africa, most of the resources are being channelled towards HIV treatment programs resulting in the neglection of hypertension, which now has higher prevalence percentages than HIV (11). Hypertension is being managed poorly with lower levels of awareness and control (11). There is sufficient evidence which shows a gap on how hypertension is being managed compared to other diseases. One of the reasons might be poor primary health care facilities. These facilities have a record of focusing on programs such as maternal and child health rather than chronic diseases (25). In most health care facilities, nursing staff are over-burdened with the management and clinical consultations of the chronic patients (3). Hypertension is being managed in less than 10% of the total population with uncontrolled hypertension (3). LHWs may ease the workload and a task-shifting approach can be considered to close the gap in hypertension management (3).

Most of the previous studies on hypertension used basic statistical methods such as linear regression models, general t-tests and general chi-squared test. All of these methods are not really suitable for cluster randomised study design because they all tend to ignore the clustering effect. In most cases, ignoring the clustering effect lead to biased results. In this study, robust statistical methods such as Multilevel regression modelling, adjusted t-test, adjusted chi squared test were used to account for the clustering effect of the study design. Also in most of the previous studies, they used the clinic link data which was collected from the clinic records which in most cases generalise the treatment decisions for a group of patients. However, in this study individual level data was used in the analysis. Egan et al. (2019), reported that individual level data analysis is more vital in that the treatment decisions is tailored to each patient.

1.4 Justification of the Study

Lay Health Workers(LHW) intervention have a positive impact on the management of patients especially those with chronic illnesses (25). The intervention might also have a positive impact on the primary care clinics in Agincourt. The intervention might bring improvement on hypertension management by improving the diagnosis of patients, patient care retention, and patient adherence to treatment (11). LHW might also facilitates task shifting where nurses in Agincourt clinics can only focus on the clinical management of patients with uncontrolled hypertension (3). Task shifting might help to control hypertension and to maintain it within patients (3). Most of the studies which were done on hypertension were focusing on European countries and mainly urban areas (1). However, this current study will be focusing on Agincourt rural South Africa (14). In this study, multilevel models will be used to assess the effectiveness of the LHW in the management of uncontrolled hypertension in Agincourt, South Africa. Multilevel models enables the simultaneous examination of group level (cluster and division) and individual level variables on individual level outcomes while accounting for the non-independence of observations within groups (34). The analysis allows the examination of between group (clinics) and within group (clinics) variability as well as how group level and individual level variables were related to variability at both levels (34).

1.5 Current Study Aim

To assess the impact of using multilevel logistic regression models with random effects to assess the effectiveness of LHW intervention in the management of uncontrolled hypertension in Agincourt, South Africa.

1.6 Primary study Aim

To assess the effect of hypertension management of lay health workers (LHW) working in South African rural primary healthcare clinics to support the provision of integrated chronic care (3).

1.7 Current Study Research Question

Does using multilevel logistic regression models with random effects to assess the effectiveness of LHWs give different results from the primary study by Goudge et al. (2018).

1.8 Primary study research question

Does LHWs intervention improve the management of hypertension by providing extra support from LHWs to the nurses who are responsible for the management of chronic conditions in primary care clinics (11).

1.9 Objectives of the study

- To describe the social demographic characteristics of the participants at baseline and post-intervention adjusting for the cluster-randomized design.
- To assess the effectiveness of the LHWs intervention given baseline and end-line differences in proportions for uncontrolled hypertension between control and intervention groups.
- To assess the effectiveness of the LHW intervention and to identify patient factors associated with uncontrolled hypertension using multilevel logistic regression models.

Chapter 2

Literature Review

2.1 Outline of the literature review chapter

The literature review starts by defining controlled and uncontrolled hypertension. It then discusses hypertension effects on a global scale which constitute the whole world and on a local scale looking at Africa and South Africa focusing more with Mpumalanga province. The role of Lay Health Workers in the management of hypertension has been discussed. Management of hypertension is key in lowering mortality and the cardiovascular diseases associated. Therefore, more detail has been discussed on the management and risk factors of hypertension. Previous methodologies that were used in different studies have been discussed in comparison to the current methodologies highlighting the strengths of the current methodologies in filling the gap left by the previous methods.

2.2 Controlled and Uncontrolled Hypertension

Controlled and uncontrolled hypertension can be defined differently. In 2007, systematic review of hypertension by Addo et al., defined controlled hypertension as systolic blood pressure (SBP) less than 140 mmHg and diastolic blood pressure (DBP) less than 90 mmHg. Uncontrolled hypertension was defined as SBP greater than or equal to 160 mmHg and DBP greater or equal to 95 mmHg or SBP greater than or equal to 140 mmHg and DBP greater than or equal to 90 mmHg (2). Singh et al. (2017), defined controlled hypertension as SBP less than 120 mmHg and DBP less than 80 mmHg (15). Uncontrolled hypertension was defined as the average blood pressure determined by SBP greater than 140 mmHg and DBP greater than 90mmHg (15). Both authors have a similar definition of uncontrolled hypertension and a slight difference on controlled hypertension. Controlled hypertension is not harmful to patients because the disease is managed. In another early review in 2009, Hurlburt et al., defined controlled hypertension as SBP less than 140 mmHg and DBP less than 90 mmHg (16). In the same study, controlled hypertension was described as SBP less than 130 mmHg and DBP less than 80 mmHg only in patients with diabetes which offers yet another definition of controlled hypertension (16). In a cross-sectional study, Masilela et al. (2020), defined controlled hypertension as SBP less than 140 mmHg and DBP less than 90 mmHg. Uncontrolled hypertension was defined as SBP greater than 140 mmHg and DBP greater than 90 mmHg. The definitions used in this current research are from the primary study.

The primary study of this research defined controlled hypertension as SBP less than 140 mmHg and DBP less than 90 mmHg (11). Uncontrolled hypertension in the primary study was defined as SBP greater than 140 mmHg and DBP greater than 90 mmHg which confirms the definition Addo, Hurlburt and Masilela findings. Therefore, in the current study, the definition of controlled and uncontrolled hypertension of the primary study, Goudge et al. (2018) will be used. High BP can be a cause of death.

2.3 Controlled and Uncontrolled Hypertension Globally

Hypertension might be one of the leading causes of death amongst all other chronic illnesses in the world today. Hansen and Taylor (2017), suggested that hypertension can be a leading preventable cause of premature death and disabilities. A systematic review of population studies, has examined that approximately 31.1% of the world's population aged 18 and above had hypertension (1). In high income countries 28.5% of the adults had hypertension and in lower and middle income countries about 31.5% had hypertension (1). In 2010, about 1.39 billion people in the whole world had hypertension. Those from high income countries were about 349 million and about 1.04 billion were from low and middle income countries (1). The global statistics for mortality from hypertension are well documented.

The systematic review of population-based studies from 90 countries detailed the statistics. The data estimated that 26.4% of the adult population, or approximately 972 million people of the world's population had hypertension in the year 2000. (4). By 2010 adult population with hypertension was expected to rise (4). Out of 1.4 billion people, only 14% had controlled SBP, which was below 140 mmHg, and DBP below 90 mmHg (4). According to the 2019 Journal of Hypertension, the projection of this global burden was expected to rise even further from 972 million to 1.6 billion by 2025. Africa has been affected severely by hypertension.

2.4 Controlled and Uncontrolled Hypertension in Africa

Hypertension is the leading cause of death amongst all other chronic illnesses on the continent of Africa. Across Africa, uncontrolled hypertension is estimated to be higher in urban areas than in rural areas (5). Refined diets and lots of starchy food often cause obesity in adult populations and hence increase hypertension (5). In of the reviews, Addo et al. (2007), the level of hypertension in Cameroon rural was about 11% of people (2). About 10% of people in urban Cameroon obtained treatment for hypertension. In Ghana by comparison, 32% of people got treatment (2). In South Africa, Addo reported that 47% of females had hypertension. The same study argued that females were more likely to be on treatment for uncontrolled hypertension as compared to males (2). Seedat et al. (2010), reported that over 500 000 people were estimated to have hypertension in South Africa (6). In a cross sectional study, Gawatudde et al. (2015), hypertension was considered a widespread problem reported in some African communities to be as high as 38% (7). Out of 650 million people, an estimate of 10 to 20 million have uncontrolled hypertension (7). The prevalence of uncontrolled hypertension in African countries is rising.

In Ethiopia, adult population with uncontrolled hypertension is approximated to be 150 million by 2025 (8). In a cross-sectional study, Tesfaye et al. (2017), emphasised controlled hypertension is very low in Africa ranging from 2.6% in Kenya to 17.8% in Namibia (8). Tesfaye et al. (2017), stressed that only few studies done in Ethiopia were in agreement with the prevalence of uncontrolled hypertension among patients on treatment. The study reported 53.4% in Gondar hospital to 59.9% in Tikur Anbessa Hospital. In a case control study, Nyaradzai et al. (2015), hypertension prevalence among patients on antiretroviral therapy (ART) in Kadoma City in Zimbabwe was estimated to be 30% (9). Approximately 61% HIV patients on antiretroviral therapy had uncontrolled hypertension, compared with 46% in the general population who had uncontrolled hypertension (9).

The hypertension levels are increasing rapidly in Sub Saharan Africa (SSA). According to an early study by Addo et al. (2007), about 25% of the people on treatment for uncontrolled hypertension seem to be well controlled (2). Across Sub Saharan Africa, it is estimated that less than a quarter of the patients reach their treatment goal (2).Goudge et al. (2018), argued that this could be due to the over burdening of health workers, in general. In SSA, it is estimated that less than 30% of the people with hypertension are on drug treatment, and less than 20% have achieved controlled hypertension (2). South Africa is believed to be amongst the countries with high prevalence of uncontrolled hypertension.

2.5 Controlled and Uncontrolled Hypertension in South Africa

South Africa is considered to have the highest prevalence of uncontrolled hypertension in Southern Africa. Jongen et al. (1983), discovered that about 25% of people from the zulu tribe had uncontrolled hypertension. The figure was considered to be the highest in urban blacks of the Zulu tribe (10). Uncontrolled hypertension in white people was around 17%, slightly lower in ethnic Indians, and 9% in blacks in rural areas (10). About 90% of the Zulu patients had undiagnosed hypertension. Approximately 58% of Indian patients and 77% of whites had hypertension that was untreated or uncontrolled (10).

According to Norman et al. (2007), figure 2.1, the impact of uncontrolled hypertension on mortality in South Africa was approximately 71.5%. The impact was more dominant in people aged 30 years and above. In that year, hypertension recorded the highest mortality rate than ischaemic heart diseases with 41.7%, stroke with 49.6% and other cardiovascular diseases (CVD) with 21.6%. Different scholars reported different hypertension levels in South Africa.

CONDITION	% OF DEATHS DUE TO HYPERTENSION
ISCHAEMIC HEART DISEASE	41.7
STROKE	49.6
HYPERTENSIVE DISEASE	71.5
OTHER CVD	21.6

Figure 2.1: **Impact of hypertension on mortality in people aged 30 and above years in South Africa in 2000**

Jongen et al. (2019), reported high uncontrolled hypertension in South Africa approximated to be 60% which is divided to be 30.4% men and 29.6% women. In a cross-sectional study, Masilela et al. (2020), uncontrolled hypertension prevalence in South Africa was 56.83%. Results from both authors are lower than the 75.5% that was reported in an earlier study by Adeniyi et al. (2016), for Eastern Cape province. Masilela et al. (2020), also found that mostly individuals aged 50 years and above had the highest levels of uncontrolled hypertension in South Africa. In his findings, almost eight out of ten people aged 50 and above had hypertension (17). Another finding from the Heart and Stroke Foundation of South Africa highlighted that at least one in three South Africans who are 15 years and older had hypertension (17). Therefore, the foundation reports compliments Masilela's findings. This study will be focusing on hypertension in Agincourt, South Africa.

Agincourt is a subdistrict of Mpumalanga province in South Africa. Uncontrolled hypertension prevalence is approximately 40% in Agincourt (11). The figure is believed to be greater since almost half of the population does not know their blood pressure status. Goudge et al. (2018) argued that this might be due to poor management of chronic illnesses in this area. In response to poor chronic illness management, lay health workers (LHW) intervention was used in managing hypertension in Agincourt.

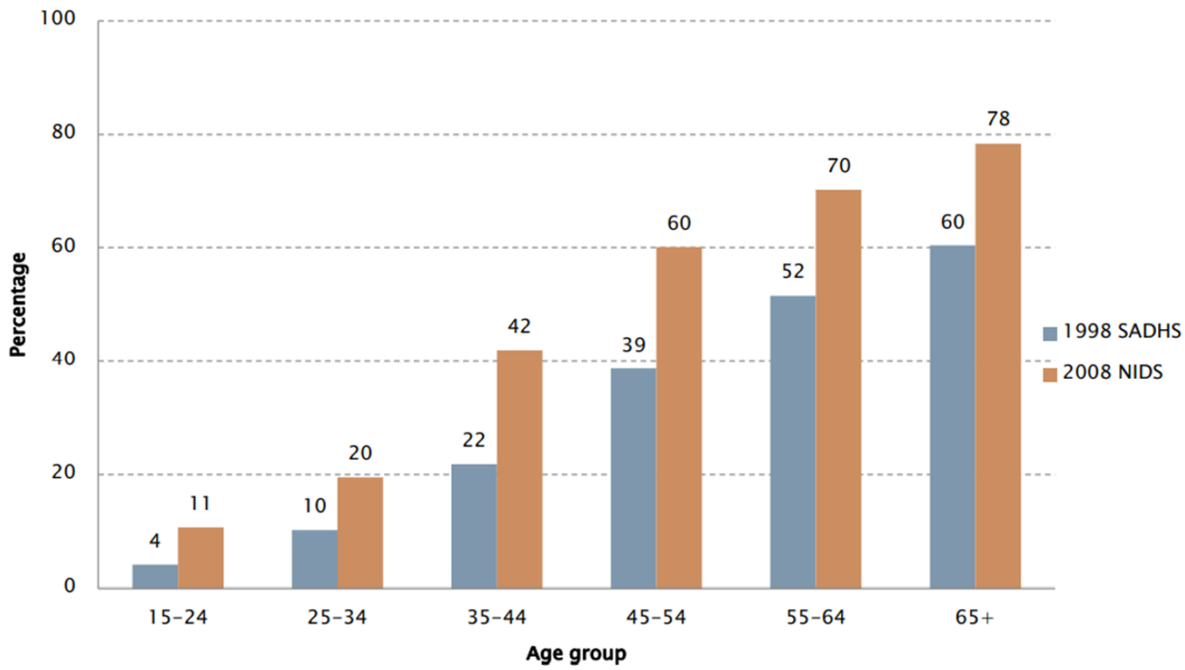


Figure 2.2: **Trends in the prevalence of uncontrolled hypertension in South African men by age group**

Bradshaw et al. (2011) highlighted that hypertension prevalence in South African men increases by age. Both in 1998 and 2008, the prevalence of hypertension in South African men was increasing by age. The elderly men 65 years and above had the highest prevalence of uncontrolled hypertension and hence men in the age group 15 to 24 had the least prevalence of uncontrolled hypertension.

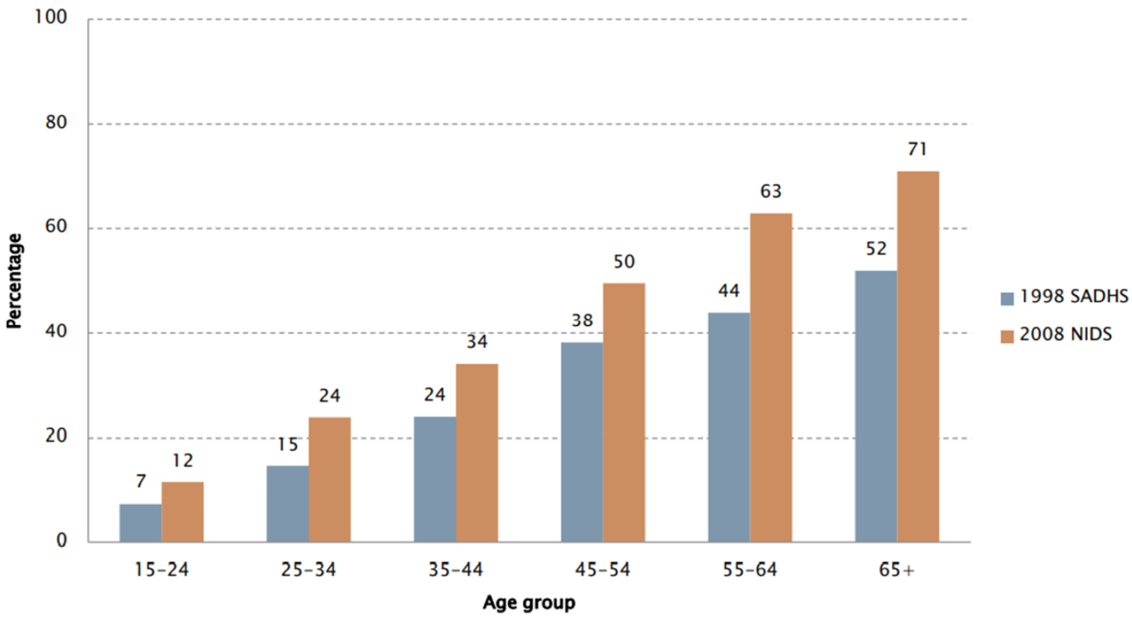


Figure 2.3: **Trends in the prevalence of uncontrolled hypertension in South Africa women by age group**

Bradshaw et al. (2011), highlighted that hypertension prevalence in South African women increases by age. Both in 1998 and 2008 the prevalence of hypertension in South African women was increasing by age. The elderly women 65 years and above has the highest prevalence of uncontrolled hypertension.Young women in the age group 15 to 24 has the least prevalence of uncontrolled hypertension.

Bradshaw et al. (2011), made a point that uncontrolled hypertension is affected by age. The graphs by Bradshaw et al. (2011), show that the hypertension prevalence in both males and females increases by age. However, according to the graphs, prevalence of uncontrolled hypertension is more higher in men than in women in South Africa.

2.6 The role of Lay Health Workers intervention in the management of hypertension

Daniels et al. (2012), defined LHWs as any health worker carrying out duties that are health related. LHWs are trained in context of the intervention, and they may not have formal professional qualifications like degrees (18). Hodgins et al. (2016), emphasised that LHWs are very important in that they can form part of interventions that aim to serve individuals and communities that might be hard to reach. They deliver health activities with the support from health professionals (13). Therefore, the difference of LHWs from other health workers is mainly on their lack of formal professional training (13).

The concept of LHWs programs began in many countries around the 1970s and 1980s, for example, in national programs such as China's Barefoot Doctors according to DiCarlo et al. (2018). The evidence of their effectiveness in the health workforce crisis has increased investments in national programs for the service they provide in many countries (11). , LHWs services are critically important in the African context. LHWs fill the gap where there is an inadequate training of qualified health professionals and understaffed health-care facilities due to poverty and emigration (11).

LHWs interventions play a vital role in health facilities. LHWs perform tasks such as management of nutrition, monitoring of maternal and child health (20). LHWs also help in improvement of primary health care, assisting malaria patients, monitoring, and assisting tuberculosis patients and HIV/AIDS prevention and control. LHW assist in psychiatric wards, and in dealing with non-communicable diseases (20).

Traditionally, controlled hypertension is assumed to be treated by medication and other medical drugs (21). In this study, LHWs intervention effectiveness will be assessed in hypertension management. DiCarlo et al. (2018) revealed that there is a rise in the interest of LHWs in the modern world. European countries are leading in utilising LHWs and Africa is also taking a positive path regarding the appreciation of LHWs (22).

2.7 The role of Lay health Workers in managing and controlling hypertension in a resource constrained environment

Lay Health workers help in improving the health of the patients in a resource constrained environment. LHWs also help connect patients who do not have access to health facilities with organizations that offers free or cheap primary care visits (35). The connection assists participants to stay updated on their hypertension condition. Tsolekile et al. (2018), highlighted that LHW offers different roles in managing and controlling hypertension.

Figure 2.4 highlighted most reported roles that LHWs perform. The distribution of medication has the highest percentage which is 84%. This is followed by advise on diet with 72%, blood pressure measurement with 63% and physical activity sessions with 53%. Only 21% were doing pill counts (35).

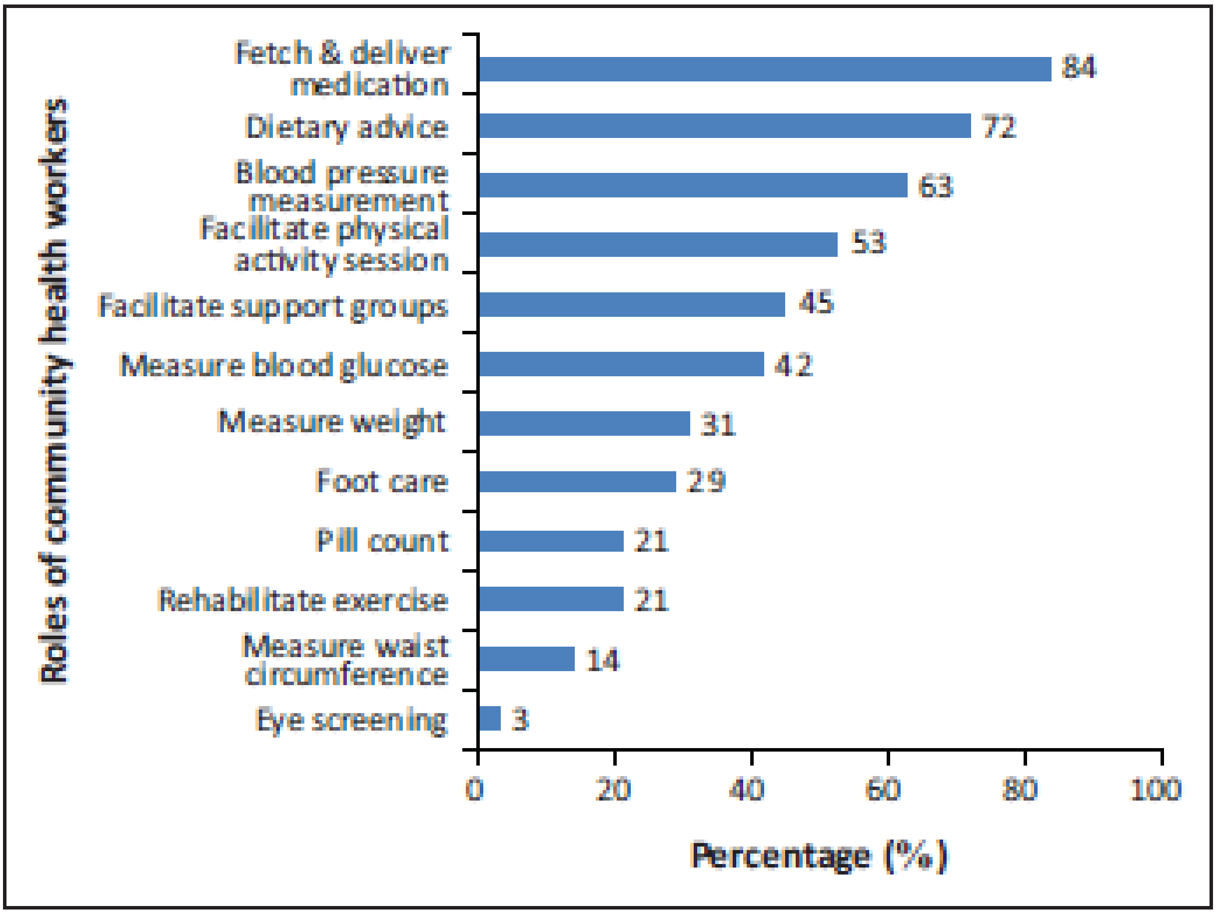


Figure 2.4: **Roles of Lay Health Workers in a resource constrained environment**

2.8 The intervention Management of Hypertension

Hypertension control is crucial in lowering the risk of cardiovascular disease (CVD), which is linked to a high rate of death and disability in adults. Controlling hypertension is essential for lowering the risk of future cardiovascular events (CVE) in a population. Controlled hypertension may be the most essential indicator for assessing the intervention overall effectiveness (23). Improving hypertension management has usually been described as increasing awareness, medication, and control. Assessing indicators' progress can help uncover potential service gaps and provide opportunity for intervention (14). In South Africa, few people are aware of their hypertension status.

A population's lack of awareness of hypertension could imply little access and involvement with the healthcare system (23). Low awareness necessitates community- and clinical-based strategies to enhance BP screening. If more people with hypertension are to be found, action is needed. Low treatment rate can be identified by lacking of access to health care, cost barriers and low rates of health literacy. Nonadherence to recommended treatment by individuals or no action by clinicians may be indicative of low treatment. In clinical setups and communities, effective interventions that offer knowledge, tools, and resources are needed in reducing treatment gaps to ensure that people receive recommended care (23). Educating people to manage hypertension is essential.

Goals and strategies can be used to control hypertension in adults due to high population rates (24). Many of the clinical and community approaches are applicable across the lifespan. Strategies and resources are required for young people's unique aspects of blood pressure management. Jackson et al. (2018), presented their findings in different publications (24). Clinicians have suggested different treatment methods.

In 2018 Ritchey et al., reported that in most cases patients with uncontrolled hypertension respond well to recommended treatments. Clinicians use proof-based interventions which help to avoid the negative impacts of uncontrolled hypertension in patients (25). Arnett et al., found in 2019 that sectors can help in adapting, executing, and extending these interventions across Africa. However, most of the times hypertension control is mostly achieved in clinical settings and communities.

Young et al. (2018), medical services providers have shown significant improvements in hypertension control. Clinicians have prioritised management of hypertension, executing evidence based procedures, committed to observing improvement after some time (26). The same study found that methods such as enhancing electronic health record systems and providing decision support tools can help in monitoring hypertension. Moreover, intensifying the use of integrated care teams might also helps. Ultimately, using endorsed shared management through self-measuring blood pressure monitoring was significant (26). Obstacles might be experienced in the management of hypertension.

Most health systems and communities have managed to contain uncontrolled hypertension, however, some difficulties still exist. Hindrances incorporate people having insufficient health care coverage inclusion, high deductibles and co-instalments (26). Lack of access to health care services also affects uncontrolled hypertension treatment. Lower levels of health literacy could prevent patients from using medications. Technology may also help in making instruments that measures blood pressure which might be of importance in monitoring uncontrolled hypertension (26). Advice from a clinician might be a crucial step that engages patients to deal with their blood pressure. Patient freedom can be inadequate when not enough upheld by strategy, foundational, and natural mediations, which might make it simpler for patients to settle on sound decisions (26). Mental challenges also affect management of hypertension and in order to obtain constructive goals, there is need for a mental health clinician. Hindrance might occur in cycle and can be either joined or eased by individuals' authentic or existing encounters, convictions, inclinations, and societies (26). Role players are vital in the management of hypertension.

In 2019, Ghosh and Kumar found that significant improvement in hypertension control will take pensive efforts by the key players. Patients can be encouraged to make use of facilities of the different sectors. The sectors include public health, clinical medicine, to mention just a few (27). Making hypertension control a national priority allows individuals from different sectors to work together to accomplish the best cardiovascular health for our nation (27).

In summary, although hypertension management has improved significantly over the years, hypertension remains a silent killer and major public health problem in Africa (27). Africa is incapacitated by poverty which according to other studies it is an underlying factor for hypertension and CVDs (28). Controlled hypertension reduces the risk of associated CVDs. The complications associated with uncontrolled hypertension include stroke, death, chest pain, heart attack, and others (29).

2.9 Risk Factors of Uncontrolled Hypertension

There are different types of risk factors of uncontrolled hypertension. In most cases, the factors can be categorised into modifiable and non-modifiable risk factors. The modifiable risk factors involves things like life style for example excessive salt consumption, a diet high in fat and acids (27). Modifiable risk factors also include lack of physical activities, too much tobacco and alcohol consumption, and low intake of fruits and vegetables (27). Non-modifiable risk factors are hereditary for example a family history of hypertension, being over the age of 65 years, and other co morbidities such as diabetes (30).

Boaten et al. (2019), stressed that age, gender (female), absolute lack of exercise, lots of salt, fast and homemade foods have a positive association with uncontrolled hypertension (31). Singh et al, (2017), also reported that age, marital status, occupation, social-economic status, education, BMI, abdominal obesity, tobacco use, alcohol use and lack of physical activities to have a positive association with uncontrolled hypertension(15). Andriolo et al. (2019), reported a list of high-risk factors such as abdominal adiposity, general adiposity, smoking, low adherence to physical activity and diet which helps stop hypertension to be positively associated with uncontrolled hypertension(32).

Different authors reported that lack of physical activities is associated with uncontrolled hypertension. However, other factors such as age, use of drugs and poor diet were also flagged to be associated with uncontrolled hypertension and medical doctors were encouraged to consider those factors as important when treating the hypertensive patients (32).

2.10 Statistical Methods Used in Previous studies

The multilevel mixed effect regression model was only mentioned in the primary study that it was intended to be used for the sensitivity analysis which was to test the difference in the change in the proportions with moderate or greater added risk of CVD in the intervention and control clusters. However, sensitivity analysis was not conducted for both primary and secondary outcome in the primary study (11). In this study, multilevel mixed effect regression models were used to assess the effectiveness of the LHWs intervention and to identify factors that are associated with uncontrolled hypertension.

Jongen et al. (2019), Basic descriptive statistics and logistic regression models were used to analyse quantitative data. One of the advantages of using a logistic regression is that it is very easy to implement and to interpret. Logistic regression models are good for simple data sets and they perform well when the data set is linearly separable. However, logistic regression models have some weaknesses such as, they are normally used to make predictions for discrete functions. They also make assumptions that the relationship between dependent and independent variables is linear. Furthermore, logistic regression models do not consider the clustering effect for hierarchical data.

John et al. (2021), Baseline SBP as well as SBP at 12 and 24 was modelled using a linear mixed-effects model with repeated measures. The analysis model included the cluster random effects for clinic, for the participants as well as the Gaussian error distribution with the identity link function (33). The linear mixed models are good in preventing false positive association which might be as a result of the population structure. They also increase power in studies without sample structure. However, linear mixed models always make assumption that the outcome variable is normally distributed. Therefore, in cases where the outcome variable is binary for example in this current study, these models cannot be used.

In the current study, multilevel regression models with random effects that use robust standard error methods were used. The robust standard errors caters for the clustering and health care effect which violates the assumption of statistical independence of the data in a randomized trial.

2.11 Summary

In the Literature review, different authors defined hypertension differently. However, Addo, Hurlburt and Goudge, defined controlled hypertension as the BP greater than 140/90mmHg. Uncontrolled hypertension was defined as BP less than 140/90mmHg. Statistics for hypertension effects were presented on a global and local scale in Africa. The use of Lay health workers in clinical activities is rising in Africa. Hypertension management was discussed and in Africa controlled hypertension levels are very low. Gelband and Duse in 2011 argued that this might be due to high rates of poverty in Africa which leads to low access to health care and poor health care services. Statistical methodologies for previous studies were discussed and, in most studies, methods used ignores the clustering effect and statistical dependency of data.

Chapter 3

Methodology

3.1 Introduction

In this chapter, the study design, study site, study population, study sampling, data collection, study outcomes, study variables, data management were described. Different statistical methods were used in coming up with the results. The chi square and t-test adjusting for clustering were used to obtain the descriptive statistics for socio-demographic and self reported health and risk variables at baseline and end of intervention. Descriptive statistics were presented using frequencies, means and standard deviations. Propensity score matching was used to reduce bias in the calculation of treatment effect of the LHW intervention in reducing uncontrolled hypertension. Lastly, factors associated with uncontrolled hypertension were determined using multilevel logistic regression models.

3.2 Primary Study Overview

3.2.1 Primary Study Design

The primary study was a pragmatic cluster randomised controlled trial with repeated cross-sectional surveys (11). The primary study randomized eight clinics of which four of the clinics received the intervention (11).

3.2.2 Primary Outcome

The primary outcome was the change in the difference of percentage of clinic users who had elevated cardiovascular risk associated with high blood pressure (BP) before and after the intervention, as measured by two cross-sectional population surveys (11).

3.2.3 Primary Study Site and Population

The primary study was conducted in Agincourt rural, a subdistrict of Mpumalanga province in South Africa (11). The estimated population considered for the primary study was about 90,000 people (52,592 older than 18 years) who lived in 15,500 households in 26 villages within Agincourt rural from 2013 up to 2015 (19). The study site is marked by the arrow on the map below. Hypertension patients aged 18 years and above and residents of Agincourt rural area were the study target population (3).

Figure 3.1: **South African Map showing Agincourt sub district of Mpumalanga**

3.2.4 Primary Study Sampling

Primary sampling was done using cluster randomized sampling and the clinics in Agincourt were considered as clusters (11). Participants considered in the two cross-sectional surveys where approximately 4000 meaning that 500 people were allocated on each of the 8 clusters (19).

3.2.5 Primary Data Collection

The data was collected by the MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt) from 2013 to 2015 using questionnaires(11). Data collected includes previous hypertension status, previous diabetes status, previous stroke incident, previous and heart failure incident. Also, data on previous angina status, heart attack history, the history on family members who once had myocardial infarctions (MI) was collected. Data on the use of primary care clinics for the past one year, stroke, smoking and the clinic they prefer most was collected (1). More patients information such as blood pressure levels for participants, random blood glucose using [model M6W]) and total cholesterol of patients was collected (1).

3.3 Current Study Design

This current research was a secondary analysis based on data generated from a cluster randomized controlled primary study design. The current study adopted the design of the primary study and its sample size. The study aimed to investigate the effectiveness of lay health workers (LHWs) intervention on the management of hypertension using robust statistical methodology.

3.4 Current Study Outcome

The primary outcome for this current study was the difference of the population with uncontrolled hypertension within control and intervention groups at baseline and end line.

3.5 Current Study Variables and their Scales

Variables for this current study included Hypertension (ordinal), Systolic and diastolic BP (average normal readings for systolic diastolic). Obesity (ordinal), blood glucose (ordinal), total cholesterol, age, gender (nominal) were also used in the study. In addition, education status (nominal), marital status, social-economic status (Kuppuswamy scale), smoking (ordinal), self-reported health and risk (ordinal) were also used in the study.

3.6 Current Study Data Management

Stata 15 was used for data cleaning, coding, and generating new variables. A do file was created to capture all the codes which were used for data cleaning. A binary variable to capture hypertension status was generated using mean systolic (SBP) and diastolic blood pressure (DBP). Patients with uncontrolled hypertension were represented by 1 and those with controlled hypertension by 0. Stata 15 was also used for labelling of variables and generating new categorical variables from the existing ones.

3.7 Missing data

Exploratory data analysis was used to determine the data structure for each variable and identify any missing values. The missing data was checked to see if whether it was missing at random (MAR) or not. Multivariate imputation by chained equations (MICE) package in R statistical software was used for imputation of the missing values (36). A sensitivity analysis using the propensity score matching on both complete case data and imputed data was used to make an inference on the data if the missing percentage was below 20%.

3.8 Statistical Analysis Plan

Objective 1: *To describe the social-demographic characteristics of the participants at baseline and post-intervention adjusting for the cluster-randomized design.*
Descriptive statistics adjusting for clustering were used to describe the demographic characteristics of the participants. The mean and standard deviation were used to describe normally distributed continuous variables. Categorical variables such as gender, were summarized using frequencies and proportions. Bivariate analysis to compare differences in proportions between intervention and control arms were done using the cluster adjusted Pearson's Chi squared test and the cluster adjusted t-test. Stata syntax used can be viewed in appendix C1.

Objective 2: *To assess the effectiveness of the LHWs intervention given baseline and end-line differences in proportions for uncontrolled hypertension between control and intervention groups.*
There were imbalances in some of the explanatory variables between patients who had uncontrolled hypertension and those who had controlled hypertension (19). Due to the imbalances, there was a higher likelihood of bias. Therefore, to address the issue, propensity scores matching was used. A logistic regression model was used with a binary outcome for hypertension and explanatory variables with imbalances to generate propensity scores. The generated propensity scores were then used to match the two groups. The propensity score kernel density graphs were used to assess matching. Stata syntax used can be viewed in appendix C2.

3.9 Propensity Score

Kurth et al. (2005), defined propensity score as the probability of receiving the treatment being compared, given the covariates under investigation. Propensity score matching reduces the imbalances between the treatment groups in observed covariates and hence decreases the bias in the resultant estimator according to Rosenbaum (1983). In this study, let the **Z** (0,1) indicate the LHW intervention, and let **X** be the vector of covariates. The propensity score $e(\mathbf{X})$ for an individual is defined as the conditional probability of intervention given covariates $\mathbf{X} : e(\mathbf{X}) = Pr(\mathbf{Z} = 1|\mathbf{X})$.

3.9.1 One to One Propensity Score Matching

In this current study, the technique was used to identify pairs of treated and untreated subjects. Only the subjects with similar propensity score values were grouped together. After forming a matched sample, the treatment effect was estimated by comparing outcomes between treated and untreated subjects. Since our outcome variable was binary, the treatment effect was described using relative risk, Austin et al.(2008), suggested.

3.9.2 Average Treatment on the Treated (ATT)

Assuming that each observation was described with a pair of random variables (Y, X), let Y be an outcome variable and X an objective characteristics. For observations in the treatment group, observations in the control group with similar X values were identified. The Y values of these matching observations from the control group were used to compute the counter-factual outcome without treatment for the observation from treatment group. An estimate for the average treatment effect was obtained as the mean of the differences between the observed values from the treatment and the matched counter-factual values from the control group.

The Average Treatment on the Treated (ATT) equation

$$ATT = \sum_i Y_i^1 = E(Y_i^0 | T = 1) \tag{3.1}$$

The Average Treatment Effect (ATE) equation

$$ATE = \frac{T = 1}{N} A\hat{T}T + \frac{T = 0}{N} \left(\sum_i E(Y_i^1 | T = o) \right) \tag{3.2}$$

Objective 3: *To assess the effectiveness of the LHW intervention and to identify patient factors associated with uncontrolled hypertension using multilevel logistic regression models.*

Multilevel models were used to identify factors that were associated with uncontrolled hypertension. The binary outcome variable was regressed with the independent variables adjusting for clustering. The independent variables were considered as factors associated with the uncontrolled hypertension if the p-values where less than 0.05. The variables with p-values greater than 0.05 were considered as factors not associated with the effectiveness of the intervention. Level 1 multilevel logistic regression model with individual and group random effects were used to identify factors associated with uncontrolled hypertension. Stata syntax used can be viewed in appendix C3.

3.10 Multi-level logistic regression models

Multilevel models were used to identify patient factors that are associated with uncontrolled hypertension in Agincourt. The advantage of using multilevel models is that they take into consideration the clustering effect (34). Multilevel models help with the simultaneous examination of group level (cluster and division) and individual level variables. This is done on individual level outcomes while accounting for the non-independence of observations within groups. This analysis allows the examination of between group (clinics) and within group (clinics) variability and how group level and individual level variables are related to variability at both levels (34).

Since our outcome variable Y_{ij} is a binary response (0 or 1) for the ith individual in the jth unit, and let π_{ij} denote the probability of having uncontrolled hypertension (i.e $Y_{ij} = 1$). The binomial distribution is characterised by two parameters: the probability of success π_{ij} and the number of trials n. For a logistic regression model, when each data item refers to an individual response with a dichotomous outcome rather than a proportion, the denominator is always equal to one. This means that we have $Y_{ij} \sim Bernoulli(1, \pi_{ij})$

It is generally defined as

$$L\{Y_{ij} = y_{ij}\} = \binom{r_{ij}}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{r_{ij} - y_{ij}} \tag{3.3}$$

Where π_{ij} represents the probability of the subject i from sub-location j who has a covariate vector x_{ij} and $y_{ij} = 1$ indicates uncontrolled hypertension and controlled hypertension if it takes value zero.

There are r_{ij} totals and $E(y_{ij}) = \mu_{ij} = r_{ij}\pi_{ij}$ and the variance is $Var(y_{ij}) = \sigma_y = r_{ij}\pi_{ij}(1 - \pi_{ij})$. Considering that a logistic regression model can be defined as follows with cluster-level random effect u_{ij},

$$logit(\pi_{ij}(y_{ij})) = x_{ij}\beta + \phi + \theta_j + \epsilon_{ij} \tag{3.4}$$

Considering the binomial mixed-effects model, for the sub clusters, $j = 1, 2, 3, \ldots .n$, the function of $y_j = (y_{j1}, y_{j2}, \ldots .., y_{jn})^T$ given a set of sub location level random effects u_{ij}; is expressed as

$$f(Y_{ij} \mid \mu_{ij}) = \prod_{i=1}^{n_j} \left[\binom{r_{ij}}{y_{ij}} \{H(\eta_{ij})\}^{y_{ij}} \{1 - H(\eta_{ij})\}^{r_{ij}-y_{ij}} \right] \tag{3.5}$$

where

$$\eta_{ij} = x_{ij}^T \beta + Z_{ij} u_{ij} \tag{3.6}$$

Then, the exponentiation and log transformation of $H(\eta_{ij})$, can be written as

$$logit(\pi_{ij}) = log(\frac{\pi_{ij}}{1 - \pi_{ij}}) = \beta_0 + \beta_1 x_{1ij} + u_j \tag{3.7}$$

Then making π_{ij} the subject of the formula

$$ln(\frac{\pi_{ij}}{1 - \pi_{ij}}) = \beta_0 + \beta_1 x_{1ij} + u_j \tag{3.8}$$

$$\frac{\pi_{ij}}{1 - \pi_{ij}} = e^{\beta_0 + \beta_1 x_{1ij} + u_j} \tag{3.9}$$

$$\pi_{ij} = (1 - \pi_{ij}) e^{\beta_0 + \beta_1 x_{1ij} + u_j} \tag{3.10}$$

$$\pi_{ij} = e^{\beta_0 + \beta_1 x_{1ij} + u_j} - \pi_{ij} e^{\beta_0 + \beta_1 x_{1ij} + u_j} \tag{3.11}$$

$$\pi_{ij} + \pi_{ij}(e^{\beta_0 + \beta_1 x_{1ij} + u_j}) = e^{\beta_0 + \beta_1 x_{1ij} + u_j} \tag{3.12}$$

$$\pi_{ij}(1+e^{\beta_0+\beta_1 x_{1ij}+u_j}) = e^{\beta_0+\beta_1 x_{1ij}+u_j} \tag{3.13}$$

where

$$\pi_{ij} = \frac{e^{(\beta_0+\beta_1 X_{ij}+u_j)}}{1+e^{(\beta_0+\beta_1 X_{ij}+u_j)}} \tag{3.14}$$

The response probability for individual i in group j. β_0 defines the model intercept; $\beta_1, \beta_2, \ldots, \beta_n$ defines the regression coefficients corresponding to the different fixed effects covariates or independent X_{ij} in the model; $U_{ij}, ij = 0, \ldots .n$ are the random effects in the model which define the clustering.

Consider η_{ij} to be the linear predictor

$$\eta_{ij} = x_{ij}^T \beta + Z_{ij} u_{ij} \tag{3.15}$$

let y_{ij} and η_{ij} be the ith individual elements of y_i and $\eta_j, i = 1, ..., \eta_j$. let $f(y_{ij}|\eta_{ij})$ be the conditional density function for the response at observation i. Because the observations are assumed to be conditionally independent, $f(.)$ with vector inputs means

$$log f(y_j|\eta_j) = \sum_{j=1}^{n_i} log(y_{ij}|\eta_{ij}) \tag{3.16}$$

The random effects u_j are assumed to be multivariate normal with mean 0 and variance $\sum$. The likelihood for cluster j is given by

$$L_j(\beta, \sum) = (2\pi)^{\frac{-q}{2}} |\sum|^{\frac{-1}{2}} \int_{R^q} f(y_j|\eta_j) exp\left(\frac{-1}{2} u_j' \sum^{-1} u_j\right) du_j \tag{3.17}$$

$$= (2\pi)^{\frac{-q}{2}} |\sum|^{\frac{-1}{2}} \int_{R^q} exp\left\{ log f(y_j|\eta_j) - \frac{1}{2} u_j' \sum^{-1} u_j \right\}$$

Where R denote the set of values on the real line and R^q is the analog in q dimensional space. The multilevel integral of equation (3.17) is generally not tractable, We make use of numerical methods to approximate the integral. The change of variable technique will be used to convert the integral into a set of nested integrals. Gaussian quadrature will be used to evaluate the integrals.

Let $u_j = Lv_j$, where v_j is a q random vector whose elements are independently standard normal variables and L is the Cholesky decomposition of $\sum, \sum = LL'$

Then $\eta_j = x_j\beta + Z_jLv_j$, and the likelihood in (3.17) becomes

$$L_j(\beta, \sum) = (2\pi)^{\frac{-q}{2}} \int_{R^q} exp\left\{logf(y_j|\eta_j) - \frac{1}{2}v_j'v_j\right\} dv_j \tag{3.18}$$

$$= (2\pi)^{\frac{-q}{2}} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} exp\left\{logf(y_j|\eta_j) - \frac{1}{2}\sum_{k=1}^{q} v_{jk}^2\right\} dv_{j1}, ..., dv_{jq}$$

Consider a q-dimensional quadrature grid containing r quadrature points in each dimension. Let $a_k = (a_{k1}, ..., a_{kq})'$ be the vector of corresponding weights. The GHQ approximation to the likelihood is

$$L_j^{GHQ}(\beta, \sum) = \sum_{k_1=1}^{r} ... \sum_{k_q=1} \left[exp\left\{logf(y_j|\eta_{jk})\right\} \Pi_{p=1}^{q} w_{k_p}\right] \tag{3.19}$$

$$= \sum_{k_1=1}^{r} ... \sum_{k_q=1} \left[exp\left\{\sum_{i=1}^{n_j} logf(y_{ij}|\eta_{ij})\right\} \Pi_{p=1}^{q} w_{k_p}\right]$$

Where $\eta_{jk} = x_j\beta + Z_jLa_k$ and eta_{ijk} is the ith element of eta_{jk}.

3.10.1 PARAMETER ESTIMATION

There are different methods of estimating multilevel parameters. The Penalized quasi-likelihood was explored since the outcome variable for this study is binary.

3.10.2 Penalized quasi-likelihood

If we consider a level 1 outcome Y_{ij} taking on a value of 1 with conditional probability π_{ij}. Then the generalized linear model or the logit model is given by,

$$ln[\frac{\pi_{ij}}{1-\pi_{ij}}] = \eta_{ij} = X_{ij}^T\gamma + Z_{ij}^T u_j \tag{3.20}$$

When looking at level one unit i nested within level two unit j, We assume variable Y_{ij} to be conditionally distributed as Bernoulli and the random effects with vector u_j to have mean zero and standard deviation one. However,in the case of binary outcomes with logit link, we will then start with the level-1 model

$$Y_{ij} = P_{ij} + e_{ij} \tag{3.21}$$

Where the $E(e_{ij}) = 0$ and $Var(e_{ij}) = p_{ij}(1-p_{ij})$ This is a non-linear model which we linearize by means of the first-order Taylor series expansion. At this iteration, we have

$$p_{ij} \approx p_{ij}^s + \frac{dp_{ij}}{d\eta_{ij}}(\eta_{ij} - \eta_{ij}^s) \tag{3.22}$$

and we evaluate at p_{ij}^s the derivative

$$\frac{dp_{ij}}{d\eta_{ij}} = p_{ij}(1-p_{ij}) = \omega_{ij} \tag{3.23}$$

Substituting the linear approximation for p_{ij} in a linear equation yields

$$Y_{ij} = p_{ij}^{s} + \omega_{ij}^{(s)}(\eta_{ij} - \eta_{ij}^{(s)}) + e_{ij} \tag{3.24}$$

Making use of algebra to rearrange all known quantities to the left-hand side of the equation result in.

$$\frac{Y_{ij} - P_{ij}^{(s)}}{\omega_{ij}^{(s)}} + \eta_{ij}^{(s)} = \eta_{ij} + \frac{e_{ij}}{\omega_{ij}}^{(s)} \tag{3.25}$$

This equation has the form of a level 2 hierarchical linear model

$$Y_{ij}^{*(s)} = X_{ij}^{T}\gamma + Z_{ij}^{T}u_j + \epsilon_{ij} \tag{3.26}$$

which gives a straightforward updating scheme. This is known as penalized quasi-likelihood because it is obtained by optimizing a quasi-likelihood (involving only 1st and 2nd derivatives) with a penalty term on the random effects. Here

$$Y_{ij}^{*(s)} = \frac{Y_{ij} - P_{ij}^{(s)}}{\omega_{ij}^{(s)}} + \eta_{ij}^{(s)}, \epsilon = \frac{e_{ij}}{\omega_{ij}^{(s)}} \sim N(0, \omega_{ij}^{(s)-1}), u_j \sim N(0, T) \tag{3.27}$$

The estimate of $\eta_{ij}^{(s)}$ can be written as below

$$\eta_{ij}^{(s)} = X_{ij}^{T}\hat{\gamma}^{(s)} + Z_{ij}^{T}u_j^{*(s)} \tag{3.28}$$

3.11 Goodness of fit

3.11.1 Intraclass correlation (ICC)

Intraclass correlations was used to test the goodness of fit for the multilevel logistic regression model. For a G-level nested random-intercept model, the g-level intraclass correlation is defined as

$$\rho^{(g)} = \frac{\sum_{l=g}^{G} \sigma_l^2}{\gamma + \sum_{l=2}^{G} \sigma_l^2} \tag{3.29}$$

The formula also applies in the presence of fixed effects covariates **X** in a random effects model. In this case, interclass correlations are conditional on fixed effects covariates and are referred to residual intraclass correlations.

Since the variance components are non negative, the ICC falls in the class [0,1]. To cater for the ICC range, logit transformation is used to obtain confidence intervals. Let $\hat{\rho}^{(g)}$ be a point estimate of the intraclass correlation and $\hat{SE}(\rho^{(g)})$ be its standard error. The $(1-\alpha) \times 100\%$ confidence interval for $logit(\rho^{(g)})$ is

$$logit(\rho^{(g)}) \pm z_{\frac{\alpha}{2}} \frac{\hat{SE}(\hat{\rho^{(g)}})}{\rho^{(g)}(1-\rho^{(g)})} \tag{3.30}$$

3.11.2 Likelihood ratio test

The likelihood ratio test was used to compare the ratio of the likelihoods of level 1 and level 2 multilevel models. The comparison was done on the log scale and the difference in log likelihood. The likelihood ratio test was used to choose the model with higher log likelihood provided that the higher likelihood was high enough.

Let L_0 and L_1 be the log-likelihood values associated with the full and constrained models, respectively. The test statistic of the likelihood-ratio test is $LR = 2(L_1 - L_0)$. If the constrained model is true, LR is approximately χ^2 distributed with $d_0 - d_1$ degrees of freedom, where d_0 and d_1 are the model degrees of freedom associated with the full and constrained models, respectively (Greene 2018, 554 - 555).

3.12 Methods Comparison

The primary study used the mixed effects model adjusting for covariates to assess the effectiveness of LHW intervention. This method is different from the multilevel logistic regression model with random effects which was used in the current study. Multilevel models with random effects were able to account for the individual differences in the response to the intervention. However, the mixed effects models only estimated the overall population level coefficient. More comparison was explored in the discussion section.

Chapter 4

Results

4.1 Introduction

In this chapter, the results for the Nkateko trial data for the period between 2013 to 2015 were presented. Descriptive statistics were presented for the control and for the intervention arms. Frequencies, means and standard deviations were presented for patients with uncontrolled hypertension for both arms. The results from the propensity score matching were presented together with the associated graphs. Multilevel logistic regression models results were presented.

4.2 Descriptive statistics

The descriptives statistics were focusing on the participants who were recorded both at baseline and at the end of the intervention. The number of participants who were recorded at both baseline and end of intervention were 810 out of 4000. This means that above 70% of the participants were recorded only once. In the control arm, about 250 participants were recorded at baseline and 213 at the end of the intervention. In the intervention arm, 185 participants were recorded at baseline and 162 at the end of the intervention. There was no much difference in control and the intervention arms on the social demographic variables at baseline and end of intervention. In the control arm, at baseline 14% of the participants had controlled hypertension and they were on treatment. At the end of the intervention, about 12.68% remained on treatment. In the intervention arm, at baseline 9.19% of the participants had controlled hypertension and they were on treatment. At the end of the intervention, there was an increase of the participants to 13.58% who were on treatment. In the control arm, at baseline, 11.20% of the participants had uncontrolled hypertension and were on treatment. At the end of the intervention, there was an increase to 19.25% of participants on treatment. This was the same case with the intervention arm, there was a rise of the participants with uncontrolled hypertension on treatment from 10.27% to 21.60%. Since most people were on treatment, there was a decrease of the participants not on treatment both at baseline and at the end of intervention in both arms. In the control arm, the participants not on treatment decreased from 31.60% to 24.41%. In the intervention arm, the participants not on treatment decreased from 31.89% to 19.14%. The proportion of participants on treatment and with controlled hypertension were higher in the control arm at baseline compared to the intervention arm. This was not the case at the end of the intervention for both arms. The proportion of patients on treatment and with controlled hypertension was higher in the intervention arm than control arm at end-line.

Table 4.1: Socio-demographic and health variables for respondents observed both at baseline and end of intervention

Variables	**control**			**Intervention**		
	baseline n=250	**end of intervention n = 213**	**P-values**	**baseline n=185**	**end of intervention n=162**	**P-values**
Mean Age (± SD)	63.58(17.26)	65.51(17.04)		61.31(17.79)	65.04(17.00)	
	n(%)	**n**(%)		**n**(%)	**n**(%)	
Gender						
Male	65(26.00)	55(25.82)	<0.001*	65(35.14)	48(29.63)	0.088
Female	185(74.00)	158(74.18)		120(64.86)	114(70.37)	
Marital status						
In a union	101(40.40)	82(38.50)	0.428	73(39.46)	57(35.19)	0.038*
Education level						
No education	131(53.04)	116(54.46)		81(43.78)	79(48.77)	
Primary	69(27.94)	51(23.94)	<0.001*	52(28.11)	42(25.93)	0.002*
Secondary	40(16.19)	42(19.72)		45(24.32)	36(22.22)	
tertiary	7(2.83)	4(1.88)		7(3.78)	5(3.09)	
Social economic status						
Upper	60(24.49)	55(25.94)		27(15.17)	28(17.39)	
Upper-middle	36(14.69)	47(22.17)		29(16.29)	27(16.77)	
Lower-middle	45(18.37)	35(16.51)	0.384	37(20.79)	36(22.36)	0.884
Upper-Lower	47(19.18)	44(20.75)		37(20.79)	38(23.60)	
Lower	57(23.27)	31(14.62)		48(26.97)	32(19.88)	
Smoking history						
Never smoked	215(86.00)	175(82.16)		148(80.00)	134(82.72)	
Previous	16(6.40)	14 (6.57)	0.002*	24(12.97)	16(9.88)	0.040*
Less than one a day	5(2.00)	8(3.76)		1(0.54)	1(0.62)	
One or more a day	14(5.60)	16(7.51)		12(6.49)	11(6.79)	
Blood pressure category						
Normal	143 (57.20)	120(56.34)		107(57.84)	96(59.26)	
Stage one hypertension	74(29.60)	63(29.58)	<0.001*	46(24.86)	34(20.99)	<0.001*
Stage two hypertension	21(8.40)	20(9.39)		20(10.81)	20(12.35)	
Stage two plus hypertension	12(4.80)	10(4.69)		12(6.49)	12(7.41)	
Health and risk						
Family history of CVD	26 (10.40)	27(12.68)		16(8.65)	10(6.17)	
Diabetes	13(5.20)	20(9.39)		14(7.57)	13(8.02)	
Coronary heart disease	15(6.00)	3(1.41)	0.019*	1(0.54)	2(1.23)	0.793
Stroke	10(4.00)	15(7.04)		5(2.70)	3(1.85)	
Heart failure	10(4.00)	6(2.82)		5(2.70)	3(1.85)	
obesity						
Male	12(18.46)	13(23.64)	<0.001*	15(23.08)	19(39.58)	<0.001*
Female	147(79.46)	137(86.71)		98(81.67)	90(78.95)	
Hypertension						
No hypertension	108(43.20)	93(43.66)		90(48.65)	74(45.68)	
On treatment and controlled	35(14.00)	27(12.68)	<0.001*	17(9.19)	22(13.58)	<0.001*
On treatment but not controlled	28(11.20)	41(19.25)		19(10.27)	35(21.60)	
Not on treatment	79(31.60)	52(24.41)		59(31.89)	31(19.14)	
Blood glucose						
Normal <11, not fasting	232(92.80)	196(92.02))		162(87.57)	152(93.83)	
High 7 <11, fasting	4(1.60)	2(0.94)		1(0.54)	1(0.62)	
Diabetic 11 or more	14(5.60)	15(7.04)		22(11.89)	9(5.56)	

* p<0.05, is significant at a 5% significance level

Table 4.1 reported that, in the control arm, there is an association between baseline and end of intervention factors such as gender (p-value <0.001), educational level (p-value <0.001), smoking (p-value = 0.002), BP(p-value <0.001), health risk factors (p-value = 0.019), and obesity (p-value <0.001). In the intervention arm, there is an association between baseline and end of intervention factors such marital status (p-value = 0.038), educational level (p-value = 0.002), smoking (p-value = 0.040), BP(p-value <0.001), and obesity (p-value <0.001).

4.3 Propensity Score matching

The logistic regression model was used in computing propensity scores. The classification or assignment of participants to groups was 61.67%. The goodness of fit test by Hosmer-Lemeshow was performed to test the significance of the logistic regression model in predicting accurate propensity scores and the results (chi2(8) = 8.95) and (P = 0.3465) were obtained. The evidence was sufficient and enough to conclude that the model was the best fit for the data in estimating propensity scores. After obtaining the propensity scores, the distribution of the propensity score was examined using the Kernel density plot in Figure 4.1.

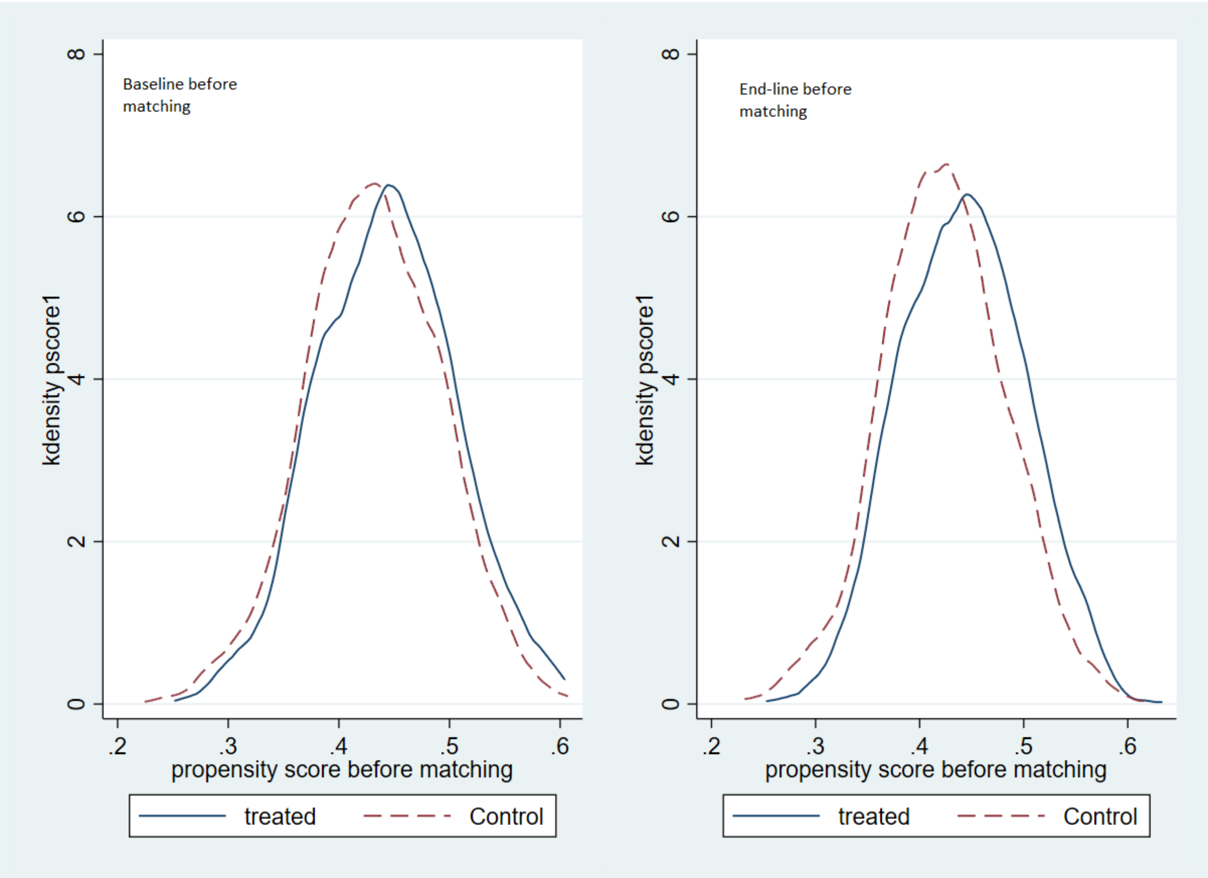


Figure 4.1: **The distribution of propensity scores before matching at baseline and at end-line respectively**

Figure 4.1 above shows the propensity scores distribution of a randomised controlled trial before matching at baseline and at end-line.These graphs shows that there was not much of an overlap as such the matching procedures should not reduce the sample sizes drastically as that would not be necessary.

4.4 Propensity score matching

Table 4.2: Average treatment effect on the treated using one to one propensity score matching

Variable	**Sample**	**Treated**	**controls**	**Difference**	**S.E**	**T-stat**
hypertension	Unmatched	0.4272	0.4324	-.0051	0.0363	-0.14
	ATT	0.4272	0.4272	0	0.0390	0.00

Table 4.2 shows the treatment effect in the treated patiences before and after matching. The average treatment effect of the unmatched was 42.72% and of the matched was 42.72%. This shows that there is no difference of the treatment effect in both matched and unmatched observations using the treatment effect scale from $0 - 100\%$.

Table 4.3: The treatment effect using inverse propensity weighting

hypertension	**Coefficient**	[95% **C-interval**]	**P>|z|**
ATE	0.0006	[-.0686, 0.0698]	0.986
ATET	0.0019	[-.0686, 0.0725]	0.957

Table 4.3 shows the average treatment effect in the whole population and in the treated population after matching the propensity scores. The treatment effect of the intervention in the population (P = 0.986) and in the treated (p-value = 0.957) is not statistically significant . Hence, in both populations groups the treatment effect was not effective in reducing hypertension levels.

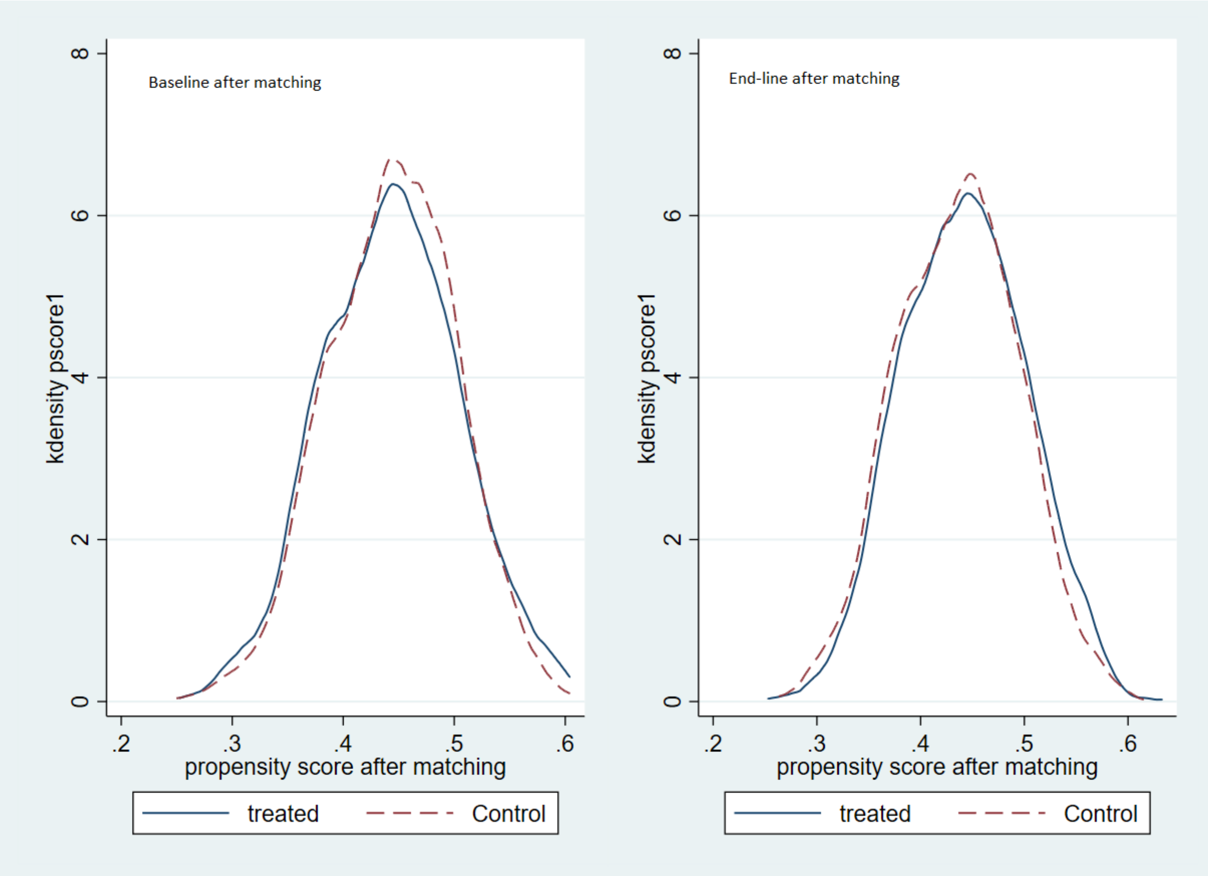


Figure 4.2: **The distribution of propensity scores After matching at baseline and end-line**

Figure 4.2 shows the distribution of the propensity scores of a randomised controlled trial after matching for baseline and end-line data. There was not much bias as the unmatched data were very similar to the matched , hence not much need of matching was necessary.

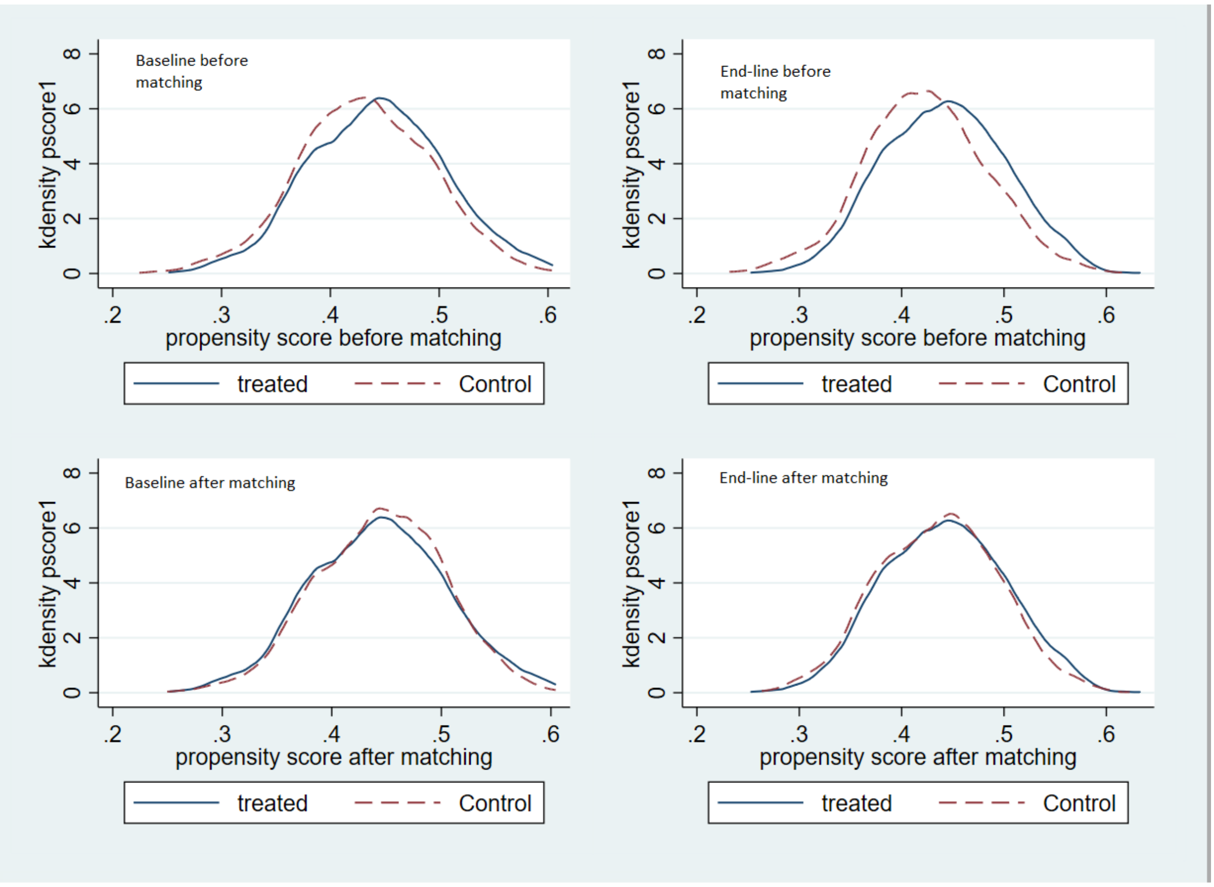


Figure 4.3: **All graphs before and after matching at baseline and at end-line**

Figure 4.3 shows all the combined graphs before and after matching. There was not much bias as the unmatched data were very similar to the matched , hence not much need of matching was necessary.

Table 4.4: Checking for bias before and after matching

Unmatched						Matched					
Variable	**Treated**	**Control**	**% Bias**	**ttest**	**P>|t|**	**Variable**	**Treated**	**Control**	**% Bias**	**ttest**	**P>|t|**
Marital status	1.625	1.605	4.200	0.600	0.552	**Marital status**	1.616	1.632	-3.200	-0.410	0.685
Gender	1.326	1.259	14.600	2.070	0.039	**Gender**	1.322	1.288	7.500	0.940	0.348
Age	63.053	64.473	-8.200	-1.150	0.249	**Age**	63.375	63.185	1.100	0.140	0.888
Education attained	0.842	0.689	17.400	2.460	0.014	**Education attained**	0.824	0.802	2.500	0.310	0.754
Total cholesterol	4.021	3.968	4.900	0.680	0.498	**Total cholesterol**	4.009	4.042	-3.100	-0.390	0.694
Obesity	1.640	1.667	-5.800	-0.820	0.414	**Obesity**	1.635	1.635	0.000	-0.000	1.000
Family risk	0.075	0.115	-13.500	-1.880	0.061	**Family risk**	0.081	0.077	1.100	0.150	0.884
Smoking history	0.326	0.315	1.300	0.180	0.857	**Smoking history**	0.316	0.338	-2.700	-0.340	0.735
Glucose level	1.184	1.138	8.600	1.230	0.220	**Glucose level**	1.099	1.093	1.200	0.190	0.852
Social economic status	3.204	2.899	21.300	2.950	0.003	**Social economic status**	3.232	3.232	0.000	-0.000	1.000
pscore	0.447	0.402	44.500	6.070	0.000	**pscore**	0.447	0.436	10.800	1.450	0.147

Table 4.4 shows that matching was effective in reducing bias in the variables. All P-values of the matched propensity scores for all the variables are not statistically significant which support randomization bias reduction.

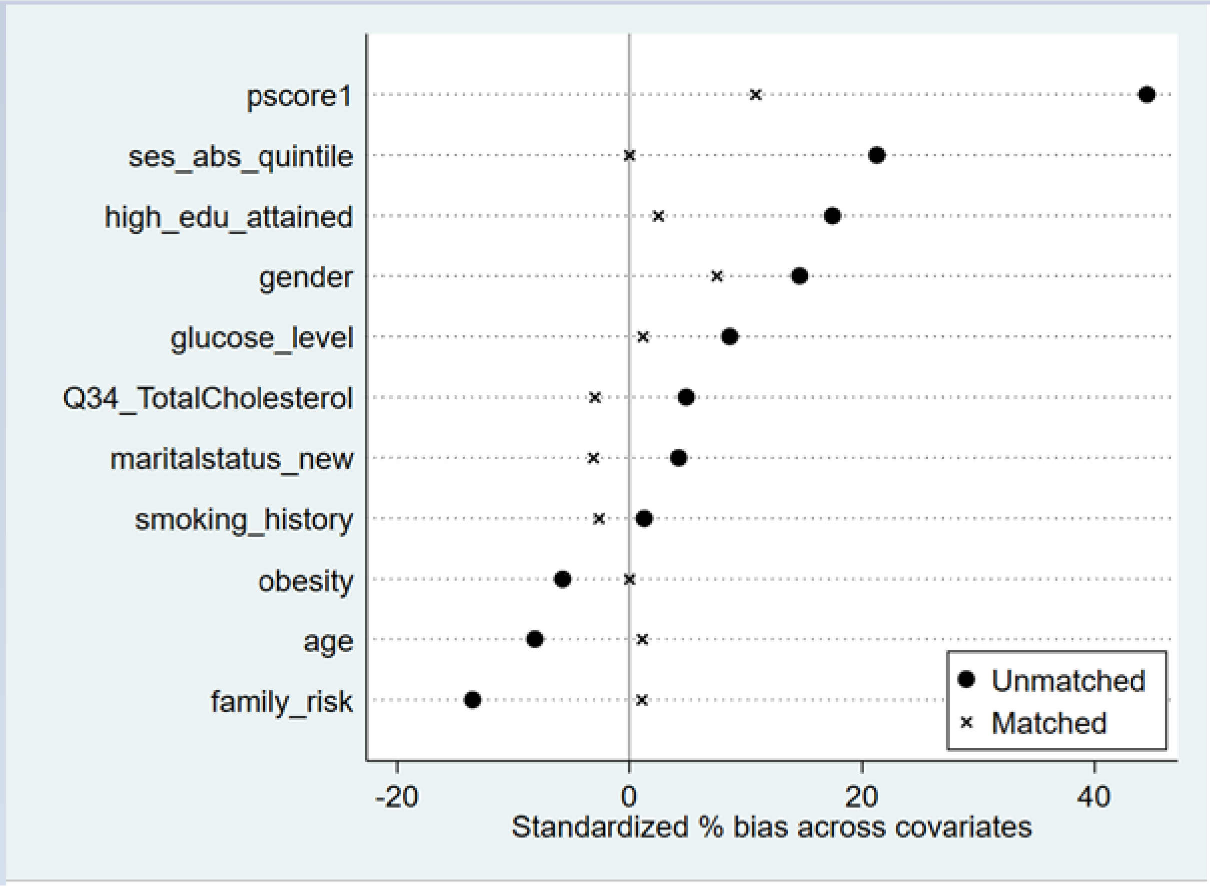


Figure 4.4: **Checking for Bias in covariates Before and after matching**

Figure 4.4 shows the bias in matched and unmatched propensity scores. There is much bias in unmatched propensity scores compared to matched propensity scores. The majority of the matched propensity scores are closer to have zero bias. Matching has a significant effect in reducing selection bias.

4.5 Multilevel logistic regression models results

Only patients who were observed at both the baseline and the end of intervention (n=810) were included. Multilevel models were used to identify patient factors associated with uncontrolled hypertension in Agincourt, South Africa. Factors such as marital status, gender, age, social economic status, total cholesterol, obesity, smoking history and family risk of cardiovascular were also included in the models. The analysis includes the use of the multilevel models, one only has individual random effects (level 1) and the other has individuals nested with a cluster random effects (level 2). The intra cluster correlation and the likelihood ratio test was used to compare the two multilevel models in order to come up with a model which best fit the data.

Table 4.5: Level one multilevel model

Covariates	Odds ratio	Std.Error	(95% C-Interval)	P-value
Fixed effects				
Maritalstatus				
In a union	Base			
Not in a union	0.8204	0.2360	(0.4668, 1.4419)	0.492
Gender				
Female	Base			
Male	2.2616	0.9699	(0.9758, 5.2414)	0.057
Age				
18-29	Base			
30-39	2.3726	0.5105	(1.5562, 3.6173)	0.607
40-49	3.7356	0.8427	(2.4007, 5.8129)	0.416
50-59	6.5870	1.6643	(4.0144, 10.8082)	0.203
60-69	6.6023	1.6921	(3.9952, 10.9108)	0.164
70-79	8.8023	2.3945	(5.1647, 15.0021)	0.148
80+	14.7532	4.4515	(8.1669, 26.6513)	0.050*
Education attained				
No Schooling	Base			
Primary	0.9403	0.3112	(0.4915, 1.7986)	0.852
Secondary	0.5040	0.2361	(0.2012, 1.2624)	0.144
Tertiary	1.9529	1.6804	(0.3617, 10.5461)	0.437
Obesity				
No	Base			
Yes	2.8629	0.9045	(1.5412, 5.3178)	0.001*
Family risk of CVD				
No	Base			
Yes	1.6526	0.6557	(0.7593, 3.5965)	0.205
Smoking history				
Never	Base			
Previous	0.2842	0.1536	(0.0985, 0.8199)	0.020*
Less than one a day	1.2334	1.0719	(0.2246, 6.7734)	0.809
One or more a day	0.5876	0.3500	(0.1828, 1.8888)	0.372
Glucose level				
Normal<11, not fasting	Base			
High 7 <11, fasting	3.7787	4.8536	(0.3048, 6.8462)	0.301
Diabetic 11 or more	0.7715	0.4551	(0.2428, 2.4516)	0.660
Social economic status				
Upper	Base			
Upper-Lower	1.0729	0.4041	(0.5128, 2.2446)	0.852
Upper-middle	1.2926	0.4817	(0.6228, 2.6835)	0.491
Upper-Lower	1.1882	0.4429	(0.5722, 2.4671)	0.644
Lower	1.4962	0.5732	(0.7062, 3.1701)	0.293
Total Cholesterol	1.1119	0.1286	(0.8865, 1.3947)	0.359
LHW intervention				
Control	Base			
Intervention	1.012	0.2428	(0.6323, 1.6196)	0.960
Intercept	0.0093	0.0157	(0.0003, 0.2554)	0.006*
Random effects				
Individual	2.9913	1.0131	(1.5402, 5.8097)	
ICC	0.4762	0.0845	(0.3189, 0.6385)	

* p<0.05, is significant at a 5% significance level

Table 4.5 reported the fixed effects and the estimated variance components of a level one multilevel model with individual random effects. From the list of fixed effects, Age (80+)(p-value = 0.050), obesity (p-value = 0.001) and previous smoking history (p-value = 0.020) were highly associated with uncontrolled hypertension. Gender (p-value = 0.057) was marginally associated with uncontrolled hypertension and the risk of males(OR = 2.2616; 95% CI: 0.9758 - 5.2414) of having uncontrolled hypertension was approximately twice that of females. The risk level of the elderly aged people (80+)(OR = 14.75; 95% CI: 8.1669 - 26.6513) of contracting uncontrolled hypertension was approximately fourteen times than that of young people aged (18 - 29). The risk level of people who were obese (OR = 2.8629; 95% CI: 1.5412 - 5.3178) of having uncontrolled hypertension was approximately three times than that of those who were not obese . The random effects showed the estimated variance components for individuals which was about 2.9913 and the related standard error was 1.0131. The intra cluster correlation (ICC) for individuals in this model was low (ICC = 0.4762; 95% CI: 0.3189 - 0.6385) which indicated a higher variability within the clusters and consequently lower variability between the clusters. The result also highlighted that the Lay health workers intervention (p-value = 0.960) was not effective in managing uncontrolled hypertension.

Table 4.6: Level two multilevel model

Covariates	Odds ratio	Std.Error	(95% C-Interval)	P-value
Fixed effects				
Maritalstatus				
In a union	Base			
Not in a union	0.7721	0.2543	(0.4049, 1.4725)	0.432
Gender				
Males	2.4702	1.2154	(0.9417, 6.4795)	0.066
Age				
18-29	Base			
30-39	1.9079	0.2978	(1.4050, 2.5908)	0.637
40-49	2.6172	0.4042	(1.9338, 3.5424)	0.444
50-59	3.8858	0.6204	(2.8417, 5.3134)	0.212
60-69	3.9197	0.6383	(2.8487, 5.3934)	0.169
70-79	4.7354	0.7821	(3.4259, 6.5454)	0.154
80+	6.8045	1.1613	(4.8700, 9.5073)	0.053
Education attained				
No schooling	Base			
Primary	0.9293	0.3487	(0.4454 1.9390)	0.845
Secondary	0.4773	0.2545	(0.1679, 1.3570)	0.165
Tertiary	2.2424	2.1946	(0.3294, 15.2679)	0.409
Obesity				
No	Base			
Yes	3.4134	1.3438	(1.5779, 7.3840)	0.002*
Family risk of CVD				
Yes	1.7676	0.8009	(0.7273, 4.2962)	0.209
Smoking history				
Previous	0.2314	0.1475	(0.0663, 0.8074)	0.022*
Less than one a day	1.4989	1.4976	(0.2115, 10.6237)	0.685
One or more a day	0.5923	0.3951	(0.1602, 2.1893)	0.432
Glucose level				
Normal<11,not fasting	Base			
High 7 <11, fasting	4.6775	6.917802	(0.2577, 8.8992)	0.297
Diabetic 11 or more	0.7231	0.4835	(0.1950, 2.6811)	0.628
Social economic status				
Upper	Base			
Upper-Lower	1.0555	0.4493	(0.4582, 2.4312)	0.899
Lower-middle	1.2637	0.5303	(0.5551, 2.8765)	0.577
Upper-Lower	1.1263	0.4772	(0.4909, 2.5842)	0.779
Lower	1.4789	0.6387	(0.6343, 3.4479)	0.365
Total Cholesterol	1.1408	0.1512	(0.8799, 1.4791)	0.320
LHW intervention				
Control	Base			
Intervention	1.0492	0.2931	(0.6068, 1.8139)	0.864
Intercept	0.0049	0.0097	(0.0001 0.2367)	0.007*
Random effects				
Individual	2.9927	1.2903	(1.2855, 6.9670)	
Individual>cluster	1.6149	1.5958	(0.2328, 11.2016)	
ICC	0.5834	0.1107	(0.3645, 0.7737)	

* p<0.05, is significant at a 5% significance level

Table 4.6 reported the fixed effects and the estimated variance components of a level two multilevel model with both individual and cluster random effects. From the Table 4.6, the fixed effects Age (p-value = 0.053), obesity (p-value = 0.002) and previous smoking history (p-value = 0.022) were associated with uncontrolled hypertension. Gender (p-value = 0.066) was marginally associated with uncontrolled hypertension and the males(OR = 2.4702; 95% CI: 0.9417 - 6.4795) were more likely prone to having uncontrolled hypertension than females. The results showed that, the elderly and those who were obese were more likely prone to contracting uncontrolled hypertension. The odds of the elderly(OR = 6.8045; 95% CI: 4.8700 - 9.5073) of having uncontrolled hypertension was very high, approximately seven times than that of the young aged (18 - 29 years). Also those who were obese(OR = 3.4134; 95% CI: 1.5779 - 7.3840) were more likely at risk of contracting uncontrolled hypertension than those who were not. The random effects showed the estimated variance components of the individual separately and those nested within clusters. The variation between individuals in clusters was (V = 1.6149) and the standard error (SE = 1.5958). The inter cluster correlation (ICC) for patients between clusters for this model was a bit high (ICC = 0.5834) which indicated a lower variability within the clusters and consequently higher variability between the clusters. This was evidence in favour of retaining the random intercepts in the model. The result also indicated that the Lay health workers intervention (p-value = 0.864) was not effective in managing uncontrolled hypertension.

Summary

Both models level one with individual random effects and level two with individuals and cluster random effects reported findings of fixed and random effects. From fixed effects of both models, age, obesity and smoking history are factors that are associated with uncontrolled hypertension. In both models, gender is marginally associated with uncontrolled hypertension. In both models risk of males in having uncontrolled hypertension is higher than that of females. The elderly, the obese, and those with a previous smoking history had a higher risk of having uncontrolled hypertension in both models. From the random effects, the individual random effects for both models were almost the same. The variation of the individuals nested within clusters (V = 1.6149) were lower than the variation between individuals across the study (V = 2.9927; 95% CI: 1.2855 - 6.9670). The intra cluster correlation (ICC) for level one model was very low indicating a higher variability within the clusters and consequently lower variability between the clusters. The intra cluster correlation (ICC) for level two model was a bit high indicating a lower variability within the clusters and consequently higher variability between the clusters..

4.6 Goodness of fit using Log Likelihood ratio test

The null hypothesis for the likelihood ratio test states that there is no significant difference between the two models (level 1 model with individual random effects and level 2 model with both individual and cluster random effects). The likelihood ratio value chi2(1) = 1.65 and a Prob > chi2 = 0.1989. The probability of the likelihood test showed that there is not enough evidence to reject the null hypothesis, Therefore, we concluded that there is no statistically significant difference between the two models. This implies that the level two model with both individual and cluster random effects is not an improvement of level one model with individual random effects only.

Chapter 5

Discussions and Recommendations

5.1 Introduction

This chapter discusses the key findings from the analysis and comparing them with findings from previous studies on some of the key results. The strengths and limitations of the models used in this current study will be discussed and conclusion will be made.

5.2 Study findings

The main aim of this study was to apply different statistical methodology from that used in the primary study by Goudge etal. (2018), in assessing the effectiveness of LHW intervention in the management of uncontrolled hypertension. The descriptive statistics for this current study reported that there were no important difference in socio demographic and health variables between the control and the intervention arm. This was the same findings made in the primary study (11). The current study found that Lay health workers intervention was not effective in reducing uncontrolled hypertension in patients. The finding was in agreement with that of Goudge et al. (2018), in the primary study. However, even though the intervention was not effective in reducing uncontrolled hypertension, lay health workers were important in some of the clinical duties such as measuring of blood pressure, booking of patient's appointments, retrieval of patient's files a day before the appointment and refiled afterwards, providing of health education and assisting nurses with pre-packing of medication (11). They were also important in reminding patients of their appointments and following up with those patients who might have missed an appointment (11). LHW are affordable and flexible and they might be of help to the South African public health especially on task shifting. In this study, multilevel logistic regression models with random effects were used to identify factors that are associated with uncontrolled hypertension. Smoking history, obesity, age, gender were found to be factors associated with uncontrolled hypertension. Comparing the results with other studies, in a cross sectional study, Aherhe et al. (2020), age, gender and obesity were also found to be associated with uncontrolled hypertension (37). In another study, Shapo et al. (2003), smoking, obesity and gender where also found to be associated with uncontrolled hypertension (38).

5.3 Strengths and Limitations of the Study

This study allowed the use of propensity score matching which played a very important role in reducing the imbalances which were there between the participants with controlled and uncontrolled hypertension. The method was not used in the primary study leaving the imbalances being present in the patients with hypertension resulting in higher chances of bias. The treatment effect of the treated obtained after matching was not significant in reducing uncontrolled hypertension in patients. In the primary study, they have also found that LHW intervention was not effective in reduction of uncontrolled hypertension in patients.

The use of multilevel models to identify factors associated with uncontrolled hypertension was important compared to mixed effects models which were used in the primary study. Multilevel models with random effects account for the individual differences in the response to the intervention compared to the mixed effects which only estimates the overall population level coefficient. Multilevel models with random effects enables the analysis of unbalanced study designs, including unequal spaced and missing data. In this study, multilevel models made it possible to estimate the within and between groups effect at the same time. Multilevel models also provided an accurate representation of the sources of variability in the data which gave more reliable test statistics, p-values and confidence intervals. The testing of the hypothesis about the variability of the treatment effects across subgroups defined by individual level covariates, through the inclusion of cross level interaction terms was made possible. Multilevel models also offered convenience, flexibility and greater statistical power during the analysis. Also, this study enhanced the modelling of possible risk factors of hypertension and symptomatic condition adjusting for other confounding factors using multilevel logistic regression models on baseline characteristics. The modelling approach used was robust in a way that it accounted for individual and cluster heterogeneity. However, multilevel models are advanced statistical technique, which requires a solid grounding in statistics. In most cases, there is misinterpretation of results from multilevel models. Another limitation is that the data set had few time points which made it difficult to use other statistical methods that take into consideration the time effect (interrupted time series analysis).

5.4 Conclusion

The propensity score matching improved the imbalances which was there between controlled and uncontrolled hypertension patients. The matching made it easy for the calculation of the treatment effect. The LHW treatment effect from a one to one matching showed that LHW intervention was not effective in reducing uncontrolled hypertension within patients. In South Africa, management of HIV by LHW is more successful compared to the management of uncontrolled hypertension. In a study by Mahomed et al.(2016), Findings were made that below 50% of the patients who attends regular clinic visits, had controlled hypertension (39). On the other hand, approximately 90% of the patients who attends regular clinic visits for HIV had a controlled CD4 count (39). This is an indication that the management of hypertension is not being prioritised in South Africa and hence the need for more attention on hypertension management. This study also showed that when analysing randomised controlled trial data, it is important to use different statistical methods to see if whether there might be any difference on the results. If the results do not change, this increases the confidence and reliability on the results. But if the results change, there will be need to make some comparisons on the methods and make use of the results from the best model. In our case, the primary study obtained the same results as this current study. Therefore, there wont be much questions or doubts about the decisions made based on the results.

Bibliography

[1] Das C Hansen KC and Tyler JK LMS. Global Disparities. *Physiol Behav.* 2017;176(3):139–48.

[2] Addo J, Smeeth L, Leon DA. Hypertension in sub-Saharan Africa: *A systematic review. Hypertension.* 2007;50(6):1012–8.doi: 10.1161/HYPERTENSION-AHA.107.093336.

[3] Limbani F, Thorogood M, Gómez-Olivé FX, Kabudula C, Goudge J. Task shifting to improve the provision of integrated chronic care: *Realist evaluation of a lay health worker intervention in rural South Africa.* BMJ Glob Heal. 2019;4(1):1–14.

[4] Egan BM, Kjeldsen SE, Grassi G, Esler M, Mancia G. The globalburden of hypertension exceeds1.4 billion people: *Should a systolic blood pressure target below 130 become the universal standard? J Hypertens.* 2019;37(6):1148–53.doi: 10.1097/HJH.0000000000002021. PMID: 30624370.

[5] Mutowo MP, Mangwiro JC, Lorgelly P, Owen A, Renzaho AMN, Mutowo MP, et al. Hypertension in Zimbabwe: A meta-analysis to quantify its burden and policy implications. *World Journal of Meta-Analysis* 2015;3(1):54-60. doi:10.13105/wjma.v3.i1.54

[6] Seedat YK. Why is control of hypertension in sub-Saharan Africa poor? *Cardiovasc Journal of Africa.* 2015;26(4):193–5.doi:105830/CVJA-2015-065

[7] Guwatudde D, Nankya-mutyoba J, Kalyesubula R, Laurence C, Adebamowo C, Ajayi I, et al. The burden of hypertension in sub-Saharan Africa: *a four-country cross sectional study. BMC Public Health.* 2015 Dec 5;15:1211. doi:10.1186/s12889-015-2546-z.

[8] Tesfaye B, Haile D, Lake B, Belachew T, Tesfaye T, Abera H. Uncontrolled hypertension and associated factors among adult hypertensive patients on follow-up at Jimma University Teaching and Specialized Hospital: cross-sectional study. *Research Reports in Clinical Cardiology.* 2017;8:21–29.https://doi.org/10.2147/RRCC.S132126

[9] City K, Magande PN, Chirundu D, Gombe NT, Mungati M, Tshimanga M. Determinants of uncontrolled hypertension among clients on anti-retroviral therapy in Kadoma City, Zimbabwe, 2016. *Clin Hypertens* 23, 14(2017). http://doi.org/10.1186/s40885-017-0070-4.

[10] Jongen VW, Lalla-Edward ST, Vos AG, Godijk NG, Tempelman H, Grobbee DE, et al. Hypertension in a rural community in South Africa: *What they know, what they think they know and what they recommend. BMC Public Health.* 2019;19(1):1–10.https://doi.org/10.1186/s12889-019-6642-3

[11] Goudge J, Chirwa T, Eldridge S, Gómez-Olivé FXF, Kabudula C, Limbani F, et al. Can lay health workers support the management of hypertension? *Findings of a cluster randomised trial in South Africa. BMJ Glob Heal.* 2018;3(1).:e000577. DOI: 10.1136/bmjgh-2017-000577.

[12] Olaniran A, Smith H, Unkels R, Bar-Zeev S, van den Broek N. Who is a community health worker? - A systematic review of definitions. *Glob Health Action [Internet].* 2017;10(1). Available from: http://dx.doi.org/10.1080/16549716.2017.1272223

[13] Hodgins F, Gnich W, Ross AJ, Sherriff A, Worlledge-andrew H. How lay health workers tailor in effective health behaviour change : a protocol for a systematic review. *Syst Rev [Internet].* 2016;1–6. Available from: http://dx.doi.org/10.1186/s13643-016-0271-z

[14] Mirzaei, M., Mirzaei, M., Bagheri, B. et al. Awareness, treatment, and control of hypertension and related factors in adult Iranian population. BMC Public Health 20, 667 (2020). https://doi.org/10.1186/s12889-020-08831-1

[15] Singh S, Shankar R, Singh GP. Prevalence and Associated Risk Factors of

Hypertension: *A Cross-Sectional Study in Urban Varanasi. Int J Hypertens.* 2017;2017:5491838. doi: 10.1155/2017/5491838. Epub 2017 Dec 3.

[16] Hurlburt ME, Houlihan SJ, Simpson SH, Cave AJ, Flook NW, , Lord CJ, Smith LL, Sternberg HH. Hypertension treatment and control rates: chart review in an academic family medicine clinic. *Can Fam Physician.* 2009 Jul;55(7):735-41.

[17] Masilela, C., Pearce, B., Ongole, J.J. et al. Cross-sectional study of prevalence and determinants of uncontrolled hypertension among South African adult residents of Mkhondo municipality. *BMC Public Health* 20, 1069 (2020). https://doi.org/10.1186/s12889-020-09174-7

[18] Daniels K, Clarke M, Ringsberg KC. Developing lay health worker policy in South Africa: a qualitative study. *Health Res Policy Sys* 10, 8 (2012).https://doi.org/10.1186/1478-4505-10-8

[19] Thorogood, M., Goudge, J., Bertram, M. et al. The Nkateko health service trial to improve hypertension management in rural South Africa: study protocol for a randomised controlled trial. Trials 15, 435 (2014). https://doi.org/10.1186/1745-6215-15-435

[20] Wilford, A., Phakathi, S., Haskins, L. et al. Exploring the care provided to mothers and children by community health workers in South Africa: missed opportunities to provide comprehensive care. BMC Public Health 18, 171 (2018). https://doi.org/10.1186/s12889-018-5056-y.

[21] Nelissen HE, Cremers AL, Okwor TJ, Kool S, van Leth F, Brewster L, et al. Pharmacy-based hypertension care employing mHealth in Lagos, Nigeria – a mixed methods feasibility study. *BMC Health Serv Res* 18, 934 (2018). https://doi.org/10.1186/s12913-018-3740-3

[22] Dicarlo A, Fayorsey R, Syengo M, Chege D, Sirengo M, Reidy W, et al. Lay health worker experiences administering a multi-level combination intervention to improve PMTCT retention. BMC Health Serv Res 18, 17 (2018). https://doi.org/10.1186/s12913-017-2825-8

[23] Gelband H, Duse A. The Surgeon General's Call to Action to Control Hypertension. South African Med J. 2011;101(Supplement 1):552–5. https://hdl.handle.net/10520/EJC67696

[24] Jackson SL, Zhang Z, Wiltz JL, Loustalot F, Ritchey MD, Goodman AB, et al. Hypertension among youths—United States, 2001–2016. Am J Transplant. 2018;18(9):2356–60.doi: 10.15585/mmwr.mm6727a2.

[25] Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Vol. 140, Circulation. 2019. 563–595.doi: 10.1161/CIR.0000000000000678.

[26] Young A, Ritchey MD, George MG, Hannan J, Wright J. Characteristics of health care practices and systems that excel in hypertension control. Prev Chronic Dis. 2018;15(6):1–7.

[27] Ghosh S, Kumar M. Prevalence and associated risk factors of hypertension among persons aged 15-49 in India: A cross-sectional study. BMJ Open. 2019;9(12):1–9. e029714. doi:10.1136/bmjopen-2019-029714

[28] Wagner KJP, Boing AF, Subramanian S V., Höfelmann DA, D'Orsi E. Effects of neighborhood socioeconomic status on blood pressure in older adults. Rev Saude Publica. 2016;50:78.

[29] Anand TN, Joseph LM, Geetha A V., Chowdhury J, Prabhakaran D, Jeemon P. Task-sharing interventions for cardiovascular risk reduction and lipid outcomes in low- and middle-income countries: A systematic review and meta-analysis. J Clin Lipidol [Internet]. 2018;12(3):626–42. Available from: https://doi.org/10.1016/j.jacl.2018.02.008

[30] Correia JC, Lachat S, Lagger G, Chappuis F, Golay A, Beran D, et al. Interventions targeting hypertension and diabetes mellitus at community and primary healthcare

level in low- And middle-income countries:a scoping review. BMC Public Health. 2019;19(1):1–20.

[31] Amponsem-Boateng C, Zhang W, Oppong TB, Opolot G, Kumi Duodu Kyere E. A cross-sectional study of risk factors and hypertension among adolescent Senior High School students. Diabetes Metab Syndr Obes. 2019;12:1173-1180. https://doi.org/10.2147/DMSO.S213552

[32] Andriolo, V., Dietrich, S., Knüppel, S. et al. Traditional risk factors for essential hypertension: analysis of their specific combinations in the EPIC-Potsdam cohort. Sci Rep 9, 1501 (2019). https://doi.org/10.1038/s41598-019-38783-5

[33] Jr JCA, Halaand B, Shirore RM, Jafar TH. Statistical analysis plan for management of hypertension and multiple risk factors to enhance cardiovascular health in Singapore: the SingHypertension pragmatic cluster randomized controlled trial. 2021;1–18. doi: 10.1186/s13063-020-05016-4.

[34] Khan M, Shaw J. Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data. J Data Sci. 2011;Vol.9(No.1):93–110.Available at SSRN: https://ssrn.com/abstract=2019344

[35] Tsolekile LP, Schneider H, Puoane T. The roles, training and knowledge of community health workers about diabetes and hypertension in Khayelitsha, Cape Town. Curationis. 2018 Mar 26;41(1):e1-e8. doi: 10.4102/curationis.v41i1.1815. PMID: 29781697; PMCID: PMC6091590.

[36] Buuren S van, Groothuis-Oudshoorn K. Mice : Multivariate Imputation by Chained Equations in r. Journal of Statistical Software [Internet]. 2011 [cited 2020 Jul 12];45(3). Available from: http://www.jstatsoft.org/v45/i03/

[37] Woldu Aberhe et al. Prevalence and factors associated with uncontrolled hypertension among adult hypertensive patients on follow-up at Northern Ethiopia, 2019: cross-sectional study. Pan African Medical Journal. 2020;36(187). 10.11604/pamj 2020.36.187.23312

[38] Shapo L, Pomerleau J, McKee M. Epidemiology of hypertension andnassociated cardiovascular risk factors in a country in transition: a population based survey in Tirana City, Albania. J Epidemiol Community Health 2003; 57(9): 734–739.

[39] Mahomed OH, Asmall S, Voce A. Sustainability of the integrated chronic disease management model at primary care clinics in South Africa. Afr J Prim Health Care Fam Med. 2016;8(1), a1248. http://dx.doi.org/10.4102/phcfm.v8i1.1248.

APPENDICES

Appendix A: Plagiarism

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

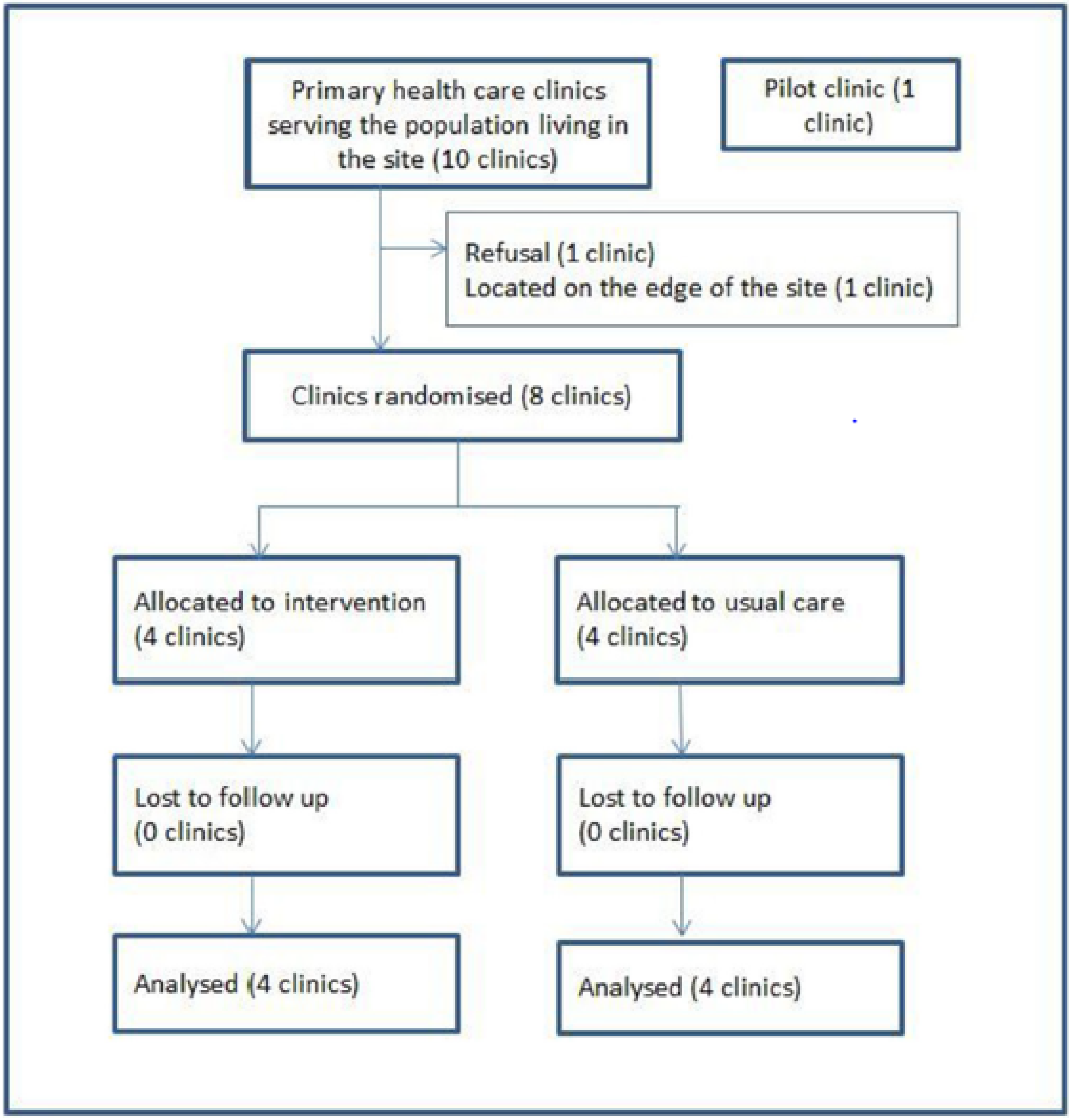
I MUNYARADZI NYAMUKOKOKO (Student number: 1816789) am a student registered for the degree of BIOSTATISTICS in the academic year 2020

I hereby declare the following:

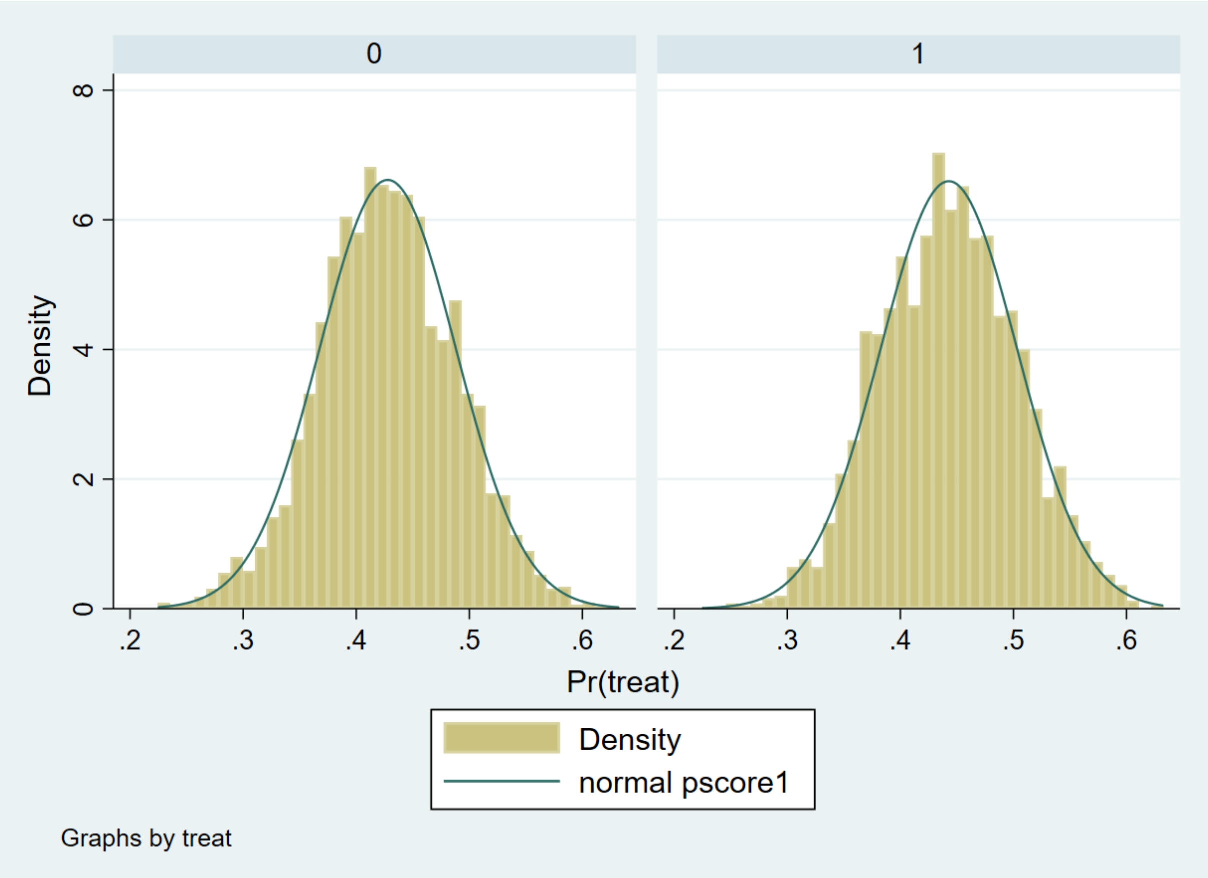
- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: ______________ Date: 28/03/2022

Appendix B: Allocation of clinics



Appendix C: Propensity score distribution



Appendix C1: Stata Code

```
*Prevalence of social demographic variables as found in two surveys
gen age = ((VisitDate- DoB)/365.25)
by UsualClinicTrialGrp, sort: summarize age if survey_round ==1
by UsualClinicTrialGrp, sort: summarize age if survey_round ==2

*Gender
tab Gender UsualClinicTrialGrp, row
tab Gender UsualClinicTrialGrp if survey_round ==1,col
tab Gender UsualClinicTrialGrp if survey_round ==2,col

*Marital status
tab marital_status
encode marital_status, gen(marital_status2)
recode marital_status2(1=2 "Not in a union")(2=1 "In a union")(3=1)(4=2)(5=1)(6=2)(7=2)(8=2), gen(maritalstatus_new)
tab maritalstatus_new
order maritalstatus_new, after(marital_status)
tab maritalstatus_new UsualClinicTrialGrp if survey_round ==1,col
tab maritalstatus_new UsualClinicTrialGrp if survey_round ==2,col

*Level of education
tab edu_code
gen high_edu_attained =.
replace high_edu_attained = 0 if edu_code =="N" | edu_code =="R" | edu_code =="C"
replace high_edu_attained = 1 if edu_code =="A" | edu_code =="B" | edu_code =="1" | edu_code =="2" | edu_code =="3" | edu_code =="4" | edu_code =="5"
replace high_edu_attained = 1 if edu_code =="M1" | edu_code =="M2" | edu_code =="M3" | edu_code =="M4" | edu_code =="M5" | edu_code =="M6" | edu_code =="M7"
replace high_edu_attained = 2 if edu_code =="M8" | edu_code =="M8" | edu_code =="M9" | edu_code =="M10" | edu_code =="M11" | edu_code =="M12" | edu_code =="N3"| ///
edu_code =="N4"
replace high_edu_attained = 2 if edu_code =="6" | edu_code =="7" | edu_code =="8" | edu_code =="9" | edu_code =="0" | edu_code =="A1" | edu_code =="A2" | ///
edu_code =="A3" | edu_code =="A4"
replace high_edu_attained = 3 if edu_code =="L1" | edu_code =="L2" | edu_code =="T1" | edu_code =="T2" | edu_code =="U1" | edu_code =="U2"
label var high_edu_attained "high education attained"
label define high_edu_attained 0 "No schooling" 1 "Primary" 2 "Secondary" 3 "Tertiary"
label val high_edu_attained high_edu_attained
tab high_edu_attained
tab high_edu_attained UsualClinicTrialGrp if survey_round ==1,col
tab high_edu_attained UsualClinicTrialGrp if survey_round ==2,col

*Social economic status
sum ses_abs_quintile
tab ses_abs_quintile
tab ses_abs_quintile UsualClinicTrialGrp if survey_round ==1,col
tab ses_abs_quintile UsualClinicTrialGrp if survey_round ==2,col
```

Appendix C2: Stata Code

```
gen Arm =.
replace Arm = 0  if arm==1
replace Arm = 1 if arm ==2


*Computing the propensity scores
logit Arm i.maritalstatus_new i.gender age i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level,vce(cluster cluster)
predict double p

histogram p,by(Arm) normal

tw (kdensity p if Arm==0) (kdensity p if Arm==1),xline(.25 .86)

*Automated propensity scores using pscores  (ado)
xi:pscore Arm i.maritalstatus_new i.gender age i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level,pscore(pscore) blockid(blockid)

*PSMATH2*
*psmatch
set seed 12345
gen x=uniform()
sort x

psmatch2 Arm,outcome(hypertension) pscore(pscore) kernel kerneltype(epan) ate
psgraph, name(gm) title(Kernel matching)
xi:pstest maritalstatus_new gender age high_edu_attained Q34_TotalCholesterol obesity family_risk smoking,outcome(hypertension) ate

*Using another matching procedure i drop the created variables
drop _pscore _hypertension

* One to one matching
psmatch2 Arm ,outcome(hypertension) pscore(pscore) odds noreplacement descending
psgraph, name(gn) title(1 - to  -1 matching)
xi:pstest i.maritalstatus_new i.gender age high_edu_attained Q34_TotalCholesterol obesity family_risk smoking,outcome(hypertension) ate

graph combine gm gn,title(Comparison of two matching techniques)
```

Appendix C3: Stata Code

```
capture log close

log using Thesis.log , replace

use NkatekoAnalyticDataset2013_2015, clear

*setting data as panel data

xtset id VisitDate

*Multilevel mixed effects models

*multilevel logistic regression model with individual random effects
melogit hypertension i.maritalstatus_new i.gender i.age_grp i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level ///
i.ses_abs_quintile arm||studyID:,or
estat icc

*multilevel logistic regression model with cluster random effects
melogit hypertension i.maritalstatus_new i.gender i.age_grp i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level ///
i.ses_abs_quintile arm ||cluster:,covariance(unstructured) or
estat icc


*Comparing multilevel regression model with level 1 and 2 random effects

drop if cluster ==.

melogit hypertension i.maritalstatus_new i.gender i.age_grp i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level ///
i.ses_abs_quintile arm ||studyID:,covariance(unstructured) or
estat icc
estimates store indiv

melogit hypertension i.maritalstatus_new i.gender i.age_grp i.high_edu_attained Q34_TotalCholesterol obesity family_risk i.smoking_history i.glucose_level ///
i.ses_abs_quintile arm ||studyID:||cluster:,covariance(unstructured) or
estat icc
estimates store full

lrtest full indiv


cap log close
```

Appendix D: Ethics clearance

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Mr Munyaradzi Nyamukokoko

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M210220

NAME: (**Principal Investigator**)	Mr Munyaradzi Nyamukokoko
DEPARTMENT:	School of Public Health
PROJECT TITLE:	Assessing the effectiveness of community health workers' intervention on the management of hypertension in Agincourt rural South Africa
DATE CONSIDERED:	26/02/2021
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Prof Eustasius Musenge and Prof Tobias Chirwa
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	10/06/2021

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions.Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report.** The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in **February** and will therefore be due in the month of **February** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date

Appendix E: Title change

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Private Bag 3 Wits, 2050
Fax: 027117172119
Tel: 02711 7172076

Reference: Mrs Sandra Benn
E-mail: sandra.benn@wits.ac.za

03 January 2022
Person No: 1816789
TAA

Mr M Nyamukokoko
zimbiru primary school
BW 800
00263
Zimbabwe

Dear Mr Munyaradzi Nyamukokoko

Master of Science in Epidemiology: Change of title of research

I am pleased to inform you that the following change in the title of your Research Report for the degree of **Master of Science in Epidemiology** has been approved:

From:	**Assessing the effectiveness of Community health workers intervention on the management of hypertension in Agincourt rural, South Africa**
To:	**Assessing the effectiveness of Lay Health Workers intervention on the management of hypertension in Agincourt rural, South Africa: A statistical methodological approach.**

Yours sincerely

Mrs Sandra Benn
Faculty Registrar
Faculty of Health Sciences

Appendix F: Turnitin Report

11%	8%	8%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	1 %
2	www.jyoungeconomist.com Internet Source	<1 %
3	Trevor A. Mori, David W. Dunstan, Valerie Burke, Kevin D. Croft, Jennifer H. Rivera, Lawrence J. Beilin, Ian B. Puddey. "Effect of dietary fish and exercise training on urinary F2-isoprostane excretion in non—insulin-dependent diabetic patients", Metabolism, 1999 Publication	<1 %
4	Shadnoush, Mahdi, Rahebeh Shaker Hosseini, Ahad Khalilnezhad, Lida Navai, Hossein Goudarzi, and Maryam Vaezjalali. "Effects of Probiotics on Gut Microbiota in Patients with Inflammatory Bowel Disease: A Double-blind, Placebo-controlled Clinical Trial", Korean Journal of Gastroenterology, 2015. Publication	<1 %
	bmcpublichealth.biomedcentral.com	