CAUSAL INFERENCE FOR LONGITUDINAL OBSERVATIONAL STUDIES WITH MULTIPLE OUTCOMES

PhD (Biostatistics) Thesis

HALIMA SUMAYYA TWABI

SCHOOL OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICAL SCIENCES

UNIVERSITY OF MALAWI

May, 2023

CAUSAL INFERENCE FOR LONGITUDINAL OBSERVATIONAL STUDIES WITH MULTIPLE OUTCOMES

PhD (Biostatistics) Thesis

By

HALIMA SUMAYYA TWABI

Master of Science(Biostatistics)- University of Malawi

Submitted to the Department of Mathematical Sciences, School of Natural and Applied Sciences, in fulfilment of the requirements for Doctorate Degree(Biostatistics).

UNIVERSITY OF MALAWI

ZOMBA, MALAWI

May, 2023

Declaration

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

HALIMA S. TWABI

FULL LEGAL NAME

_____

Signature

_____

Date

Certification

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

Signature:———————————Date: 3 May 2023

Samuel O. M. Manda, PhD (Professor)

Main Supervisor

Signature:———————————Date: 5/3/2023

Dylan S. Small, PhD (Professor)

Co-Supervisor

Signature:———————————Date: 3 May 2023

Hans-Peter Kohler, PhD (Professor)

Co-Supervisor

Signature:———————————Date:———————————

Jupiter Simbeye, PhD (Senior Lecturer)

Programme Coordinator

Dedication

To my mum and late Aunt Zahra Elizabeth Milanzi

Acknowledgements

All praise belongs to Allah for granting me the opportunity to further my education and for giving me the strength and knowledge to complete this study. First of all, I would like to thank my supervisor, Prof. Samuel Manda, who has been an immeasurable mentor, advisor, and advocate. I cannot thank you enough for the support, guidance, and learning opportunities you have provided me with throughout my work. I also wish to thank my co-supervisors, Prof. Dylan Small, for his additional wisdom, insights, and advice into my Ph.D. work and Prof. Hans-Peter Kohler's wisdom in contextualizing my Ph.D. work and applying it to the Malawi Longitudinal Study of Families and Health (MLSFH). I am greatly honored to have had an incredible team of advisors. I also want to thank my sponsors DELTAS Africa Initiative, Sub-Saharan Africa Consortium for Advanced Biostatistics Training (SSACAB) Phase I, and L'oreal UNESCO for women in Science for the tremendous support and funding. The financial support and recognition made it possible to participate in several activities that contributed to my career growth. Last but not least, I would like to thank the South African Medical Research Council for supporting some of my Ph.D. research visits to South Africa for my face-to-face meetings with Prof. Samuel Manda.

I wish to express my heartfelt gratitude to my husband, Dr. Shaffi Mdala, for his tremendous support throughout this Ph.D. journey. Your patience, constant encouragement, and, most importantly, unconditional love throughout this journey have been

instrumental in my completion. Thank you, mum and dad, Mariam and Hamisi Twabi, for your continuous prayers and support and for believing in my capabilities. To my boys, Zayd and Zubayr, thank you for being patient with me when mummy's availability was very scarce. To my mother-in-law, Khadija Mdala, my brothers, Hussein, Ahmad, Imran, and my sisters, Karida, Olivetta, and Amina, and not forgetting the boy's nanny Fatima thank you for the support and for filling my place when I was not present.

Lastly, to all my friends that have supported me along the way: Amina Ghana and Marriam Mimie James Mwalwenje, who would always lift my spirits and give me the urge to keep moving. Fiskani and Jesca were always ready for a pep talk. Chikondi and Khadija kept me sane with company and friendship. Finally, to my SSACAB family, for always being a hand away when I sort your assistance and many more who gave me their friendship and love.

Abstract

In the presence of time-varying treatment exposures and confounders, Marginal structural models with inverse probability treatment weights (MSM-IPTWs) are used to control for confounder variable imbalances in the estimation of causal effects in longitudinal observational studies. However, these methods have largely been developed for univariate outcomes. This Ph.D. thesis develops causal inference methods that could be used in longitudinal observational studies with multiple outcomes. The proposed methods are built around MSM-IPTWs and encompass two additional elements namely; a) re-defining the weights as a product of inverse weights at each time point and b) accounting for the possible correlation between and within (serial) the multiple outcomes. The capabilities of the proposed methods are demonstrated with an application to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners using the Malawi Longitudinal Study of Families and Health (MLSFH) data. The results have shown that the causal estimate of HIV positivity awareness and their associated standard errors were slightly attenuated using our proposed methods compared to the standard bivariate (condom use and multiple sexual partners) models. We recommend the use of the proposed MSM-IPTWs when estimating the causal effects of treatment exposures on multiple outcomes in longitudinal observational studies. The methods correctly control for time-varying treatment and confounders and adjust for possible dependency within and between the outcomes.

Published papers and under review

Twabi, H. S., Manda, S. O., & Small, D. S. (2020). Assessing the effects of maternal HIV infection on pregnancy outcomes using cross-sectional data in Malawi. *BMC Public Health*, 20(1). https://doi.org/10.1186/s12889-020-09046-0.

Twabi, H. S., Manda, S., & Small, D. S. (2021). Evaluating the Effect of Appropriate Complementary Feeding Practices on Child Growth in Malawi Using Cross-Sectional Data: An Application of Propensity Score Matching. *Frontiers in nutrition*, 8(1), 714232. https://doi.org/10.3389/fnut.2021.714232

Twabi, H. S., Manda, S. O. M., & Small, D. S. (2022). Estimating Unhealthy Food Effects on Childhood Overweight in Malawi Using an Observational Study. *Maternal and Child Health Journal*, 26(11), 1-9. http://doi.org/ 10.1007/s10995-022-03505-3.

Twabi, H. S., Manda, S. O., Small, D. S., & Kohler, H. P. (2022). A marginal structural model for estimation of the effect of HIV positivity awareness on risky sexual behavior. *Biostatistics & Epidemiology*, 6(2), 309-326.

Twabi, H. S., Manda, S. O., Small, D., & Kohler, H. (2023a). A comparative simulation study of multivariate regression models for multiple outcomes. *Journal of Statistical Computation and Simulation*. Under review. Submission ID: 231287550.

Twabi, H. S., Manda, S. O., Small, D., & Kohler, H. (2023b). Derivation of a multivariate longitudinal causal effects model. *Journal of Applied Statistics*. Under review. Submission ID: 230750193.

Conferences

Twabi, H. S., Manda, S. O., Small, D. & Kohler, H.P (2022, December 17-19). A Marginal Structural Model for Longitudinal Observational data with multiple outcomes. 15th International Conference of ERCIM WG on Computational and Methodological Statistics (CMStatistics), Kings College, London.

Twabi, H. S., Manda, S. O., Small, D. & Kohler, H.P. (2021, December 1-3). Conference proceedings: Marginal Structural Models to Estimate the Effect of HIV Status Awareness on HIV Risky Sexual Behaviours and Marriage dissolution. SASA Conference presentation (Virtual).

Twabi, H. S., Manda, S. O., & Small, D (2018, July 8-13). Assessing Causal Effect of HIV Knowledge on Sexual Behaviours and Marriage Outcomes using Longitudinal Observational Study in Malawi. International Biometric Society Conference, Barcelona, Spain (Poster presentation).

Twabi, H. S., Manda, S. O., & Small, D (2018, September 10-14). The Effect of HIV Knowledge on Risky Sexual Behaviours using Marginal Structural Models. DELTAS Africa SSACAB AGM and Research Conference, Nairobi, Kenya (Oral presentation).

Twabi, H. S., Manda, S. O., & Small, D. (2019, September 8-11) Causal Inference for Multiple Outcomes for Observational Studies. Joint SUSAN of the IBS and DELTAS Africa Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) in Cape Town, South Africa. (Oral presentation).

Twabi, H. S., Manda, S. O., & Small, D. (2020, October 19-23). Estimating the Effectiveness of Exclusive Breastfeeding on Multiple Child Growth Measures from Observational Data in Malawi. SSACAB Virtual Scientific Conference.

Twabi, H. S., Manda, S. O., & Small, D. (2021, November 22-24). A Comparative Performance of Statistical Methods for Treatment Effect on Multiple Outcomes. A Simulation Study and Application to Child Growth Measures. SAMSA Conference presentation (Virtual).

Twabi, H. S., Manda, S. O., & Small, D. (2021, June 18). Central Botswana Mathematics and Statistics Conference, Botswana International University of Science and Technology (Virtual).

Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ANCOVA	Analysis of Covariance
ATE	Average Treatment Effect
ATT	Average Treatment Effect on Treated
DAG	Direct Acyclic Graphs
DHS	Demographic and Health Survey
GLMM	Generalized Linear Mixed Model
HIV	Human Immuno-deficiency Virus
IPW	Inverse Probability Weighting
IPTW	Inverse Probability Treatment Weights
MDHS	Malawi Demographic and Health Survey
MLSFH	Malawi Longitudinal Study of Families and Health
MSM	Marginal Structural Models
NN	Nearest Neighbour
NSO	National Statistics Office
PS	Propensity Scores
PSM	Propensity Score Methods
RCT	Randomized Control Trials
SITA	Strongly ignorable treatment assignment
SSA	Sub-Saharan Africa
SUTVA	Stable Unit Treatment Value Assumption
ZDHS	Zambia Demographic and Health Survey

CHAPTER 1 : Introduction

1.1 Estimation of treatment effect in medical research

In health research, interest is to estimate treatment exposure effects on an outcome. For example, in child growth studies, the interest would be to see the impact of a child feeding intervention on a positive change in child growth. Here we consider complementary feeding, defined as introducing solid foods to a child at the age of six months, as the treatment or exposure. A health researcher would be interested in answering the following question, "what is the effect of giving appropriate complementary foods (timely introduction of solid foods and feeding a child the recommended number of times, with the minimum dietary requirements) to a child from 6 to 23 months on change in the child's growth at five years of age". To ascertain the effect, a child would have a particular growth measurement (i.e., height) at five years of age if given appropriate complementary foods and a different height if not given suitable complementary foods. To measure the causal effect of proper complementary feeding on the child's growth, we would need to compare the outcome for the same child in both scenarios. However, since it is not possible to observe both outcomes for the child at the same time, then an individual causal effect cannot be observed, and this is known as the "fundamental missing problem of causal inference" (Holland, 1986).

One way to observe the causal effect is to take the difference of the averages between those in the treated population (or sample) and those not treated known as the average treatment effect (ATE). The main goal is to ensure that the effect obtained is mainly due to the difference in the treatment groups without interference from other factors.

Assigning subjects (for example, children) to either a treatment or a control group can be done using randomization. One has control over the assignment of the subjects within either the treated or the control group. For example, one simple randomization method is tossing a coin to determine which feeding group to allocate a child. There is a chance that through randomization, a large number of children may be assigned to the treatment group or vice versa. Through randomization, treatment assignment is independent of measured and unmeasured factors known as confounders influencing the treatment and affecting the outcome (Rosenbaum, 2002). Thus, on average, the treatment groups are similar, and there is no need to balance the two treatment groups. The difference in the growth measurements for children who were five years old can be attributed to appropriate complementary feeding assignments, which would be the only difference between the children in these two groups. Even though randomization ensures balance, studies have shown that it does not guarantee complete balance, especially for small samples (Suresh, 2011; Saint-Mont, 2015). Several randomization procedures such as stratified randomized control trial (RCT) and statistical approaches have been developed and implemented to solve this problem (Kendall, 2003).

RCT designs are considered as the gold standard for treatment assignment and to assess the exposure effect (Rubin, 1974). However, most biomedical or public health research problems may consider randomizing subjects within treatment groups as unethical and unfeasible. For example, evidence has shown the effectiveness of complementary feeding on improving child growth measurements (Twabi, Manda, & Small, 2021; Kumar, Goel, Mittal, & Misra, 2006; Walters, Rakotomanana, Komakech, & Stoecker, 2019). Therefore, randomizing children into either appropriate or not appropriate complementary feeding groups would be unethical and possibly harmful to those assigned to the group without

appropriate feeding. Thus, in the absence of an RCT, readily available observational data from demographic and health surveys, routine health information systems, electronic medical records, or disease surveillance registries which contain rich and valuable data provide opportunities to assess treatment effects. However, subjects are not randomly assigned to treatment groups for observational studies. The researcher can give an individual treatment, for example, in health research studies, where a patient has been prescribed a particular treatment after being diagnosed with an illness. However, in this scenario, the treatment allocation is not random, and the researcher then observes the outcome. On the contrary, another example falls under health survey research. The exposure and the outcome are measured simultaneously, and there is no randomization and treatment assignment. These two scenarios of the observational studies may result in selection bias.

Apart from the problem of selection bias from observational studies, treatment assignment is not independent of confounders. For the child growth example, children from wealthy families (or urban areas) would be given appropriate foods than those from poor households (or rural areas). Additionally, educated mothers would understand the importance of providing a child with appropriate foods and may have more income to afford such types of foods. Therefore, the treatment effect estimation from non-randomized studies cannot be done by comparing the outcomes between the treatment groups due to these confounding factors, which results in confounder bias. These factors include characteristics of an individual collected during a survey that may influence the causal association and distort the true causal effect. A lot of work has been done in developing methods that assist in assessing the causal association from observational studies based on the potential outcome framework (Rosenbaum & Rubin, 1984; Robins, 1986; Rubin, 1974, 2006). This Ph.D.

thesis contributes to the research area of causal inference, where the statistical analysis objective is the estimation of a treatment or exposure on an outcome. In particular, this thesis deals with estimating causal effects in an observational study where it is essential to control for covariates. We use the terms treatment and exposure interchangeably.

1.2 Estimation of treatment effects

For studies that follow an RCT design, estimation of the treatment effect can be done directly on the data with the assumption that all measured and unmeasured confounders are balanced by design. The effect is obtained using standard statistical analysis such as the t-tests for continuous outcomes, Chi-square tests for categorical data, and regression methods. Some of the traditional statistical methods used to estimate the effects from RCTs include Multivariate regression analysis, such as the analysis of covariance (ANCOVA). The ANCOVA are used to control confounding by covariate adjustment, ensuring all potential confounders are included in the regression model. In addition, the ANCOVA model are used to estimate the causal effect by removing the influence of covariates that are not balanced at randomization on the error term (Egger, Coleman, Ward, Reading, & Williams, 1985; Winkens, van Breukelen, Schouten, & Berger, 2007; Egbewale, Lewis, & Sim, 2014). The ANCOVA model have estimated the treatment effects under several assumptions. These include assuming that the treatment and regression effects are additive. In addition, the residuals must be independent and identically distributed as normal with mean zero and constant variance, and the model must be correctly specified. However, the limitation of the regression analysis in the control of confounders is that a wrong assumption about the form of the relationship between confounder and outcome can lead to incorrect

conclusions about the exposure effects (Pourhoseingholi, Baghestani, & Vahedi, 2012).

For observational studies, the researcher merely observes both the exposed and unexposed subjects. The subject's presence within a particular treatment group may be influenced by purported measured and unmeasured confounders. Therefore, the causal effect estimates using standard methods from such data may not be accurate and mere associations would be obtained. To correct the limitations encountered using standard statistical methods, statistical approaches that ensure comparability between the treatment groups have been developed. Some of these methods control for the imbalances in confounder distribution between treatment groups (Rubin, 2006; Abadie, Drukker, Herr, & Imbens, 2004) through matching. For example, given a set of observed covariates for a child consisting of age, gender, and place of residence. A researcher would pair an unexposed child with an exposed child, having the same gender, same (or similar) age, and from the same place of residence. However, as the dimensionality of the covariates increases, matching subjects concerning a large number of covariates tends to be complicated. Other methods include stratifying the exposure groups and comparing subjects within the strata. The limitations of stratification occur just like for matching with increased covariates, which may result in numerous strata, thus, having some fewer subjects within the treatment strata combination which may result in incomplete balance and estimation of the treatment effects within the strata may be biased. Due to the dimensionality problem with the standard matching and stratification, Rosenbaum and Rubin (1983) developed a propensity score, which is defined as the probability of a subject (e.g., a patient) being treated (assigned to treatment) conditional on observed confounders. Under certain assumptions, known as a strongly ignorable assumption (no unmeasured confunders present and all subjects have a non-zero

probability of being treated), Rosenbaum and Rubin (1983) showed that the propensity score is a balancing score, and conditioning on the propensity score resembles an RCT design and ensures a balance between the treated and untreated groups. Generation of the propensity score involves estimating the treatment as a function of the purported confounders from the data. The propensity score is then used to balance the confounders across the treatment groups using propensity score matching, inverse propensity score weighting, stratification, and covariate adjustment (Austin, 2011; Austin & Stuart, 2015; McCaffrey, Ridgeway, & Morral, 2004). Another approach that allows unbiased estimation of the exposure effect from observational data includes the instrumental variable approach (Hernán & Robins, 2020). The instrumental variable method has been commonly used to estimate causal treatment effects in the presence of measured and unmeasured covariates.

1.3 Causal Inference in Observational Data

Causal inference from observational studies involves removing the effect of a set of known purported confounders on the impact of a treatment on an outcome. Several assumptions and techniques are implored to achieve causal estimation from observational data. This section provides a statistical description and theoretical background of estimating treatment effects from observational data.

1.3.1 Notation

Following Rubin (1974)'s notation, suppose there are n subjects, denoted by $i = 1, \ldots, n$. Let T_i denote an observed binary exposure for subject i, $T_i = \{0, 1\}$. For the child growth

example, $T_i = 1$ if a child is given appropriate foods, $T_i = 0$ if not given appropriate foods. Let us assume that for each subject, we observe an outcome Y_i that could be either measured on a continuous or binary scale. Furthermore, let $\mathbf{X}_i$ be a vector of p measured covariates, a set of other baseline covariates or characteristics of subject i, such as demographic characteristics. To characterize the causal effect, we define potential outcomes as illustrated by Rubin (1974) in the subsequent section.

1.3.2 Framework of Potential outcomes

For a binary exposure $T_i = \{0, 1\}$, we define a set of potential outcomes Y_T, for subject i, where Y_1 is the outcome observed if $T = 1$ and Y_0 is the observed outcome if $T = 0$. The causal effect for subject i is the difference between the potential outcomes.

$$\Delta_i = Y_{1i} - Y_{0i} \tag{1.1}$$

For $i = 1, \ldots n$ and without loss of generality we generalize Y_i to be Y. For subject i, only one of the potential outcomes can be observed. Thus, the observed outcome is given as;

$$Y = TY_1 + (1 - T)Y_0 \tag{1.2}$$

Equation (1.2) illustrates the *fundamental problem of causal inference*. The fundamental problem of causal inference specifies that only one outcome is observed for a subject under the actual treatment received. Equation (1.2) is also referred to as the consistency assumption (Rosenbaum & Rubin, 1983; Austin & Mamdani, 2006). Estimation of subject-specific causal effects is not achievable due to the fundamental problem of causal inference

(Holland, 1986).

1.3.3 The Average Treatment Effect

Since, in practice, we cannot observe both potential outcomes, the interest would be to estimate the average of the individual effects across populations (or sub-populations) of interest among the treatment groups and take the difference. The average treatment effect (ATE) is given as

$$\widehat{ATE} = E[Y_1 - Y_0]$$

$$\widehat{ATE} = E[Y_1] - E[Y_0] \tag{1.3}$$

This is the average treatment effect estimated at the level of a population. Another interest would be to estimate the population average treatment effect only among treated subjects given as

$$\widehat{ATT} = E[Y_1 - Y_0|T = 1] \tag{1.4}$$

Based on equations 1.3 and 1.4 the treatment effect is estimated from the population of the treatment groups when using observational data. Instead of individual comparisons, we could compare potential outcomes across subjects in a defined population according to the treatment level within that population (Rosenbaum & Rubin, 1983; McCaffrey et al., 2004).

Another causal estimand of interest is the average treatment on the treated (ATT) given

in equation (1.4). The ATT considers a comparison of outcomes among the treated population. The choice of the causal estimand depends on the research question of interest. This chapter focuses on estimating the average treatment effect for binary and continuous outcomes.

In observational studies, given that we observe subjects in a population who are in the treatment group ($T = 1$) and another set of subjects in the control group ($T = 0$), then the comparison of the mean difference between the treatment groups would not be equivalent to the average causal effect ($\Delta\mu \neq \widehat{ATE}$). There is no random allocation of subjects into the treatment groups. Instead, the subject selects the preferred group, or a researcher observes the subject's exposure. Thus, some underlying factors, characterized by the covariate vector $\mathbf{X}$, may determine which group the subject belongs to, and the differences between the exposure groups are likely influenced by the covariates ($X's$).

1.4 Causal Inference Assumptions

1.4.1 Unconfoundedness

Identification of the causal effect in equation 1.3 is possible if we are able to estimate the mean potential outcomes from observed data $O = (Y, T, X)$. Under randomized assignment of subjects into treatment groups, the potential outcomes and treatment assignment are independent $Y_1, Y_0 \amalg T$, hence, for RCT, $E[Y|T = 1] = E[Y_1|T = 1] = E[Y_1]$. In observational studies $(Y_0, Y_1) \amalg T$ is unlikely to hold.

$$E[Y|T = 1] = E[Y_1|T = 1] \neq E[Y_1]$$

However, it is possible to identify confounding factors for the subjects. For example, consider a set of confounders in $\mathbf{X}$ in which subjects share the same value of $\mathbf{X}$. Thus, there would be no association between the treatment groups and the potential outcomes. Exposure among subjects with similar values of the confounders would be similar to a random allocation. The unconfoundedness assumption is given as;

$$(Y_1, Y_0 \amalg T|\mathbf{X}) \tag{1.5}$$

The assumption states that the potential outcomes are independent of the treatment given the confounders. This assumption further implies the absence of unmeasured confounders. Being in either treatment group does not depend on the potential outcomes.

1.4.2 Positivity (Overlapping)

Another assumption states that for every value of the confounders, $\mathbf{X}$, every subject has a probability of receiving the treatment also known as the positivity assumption. The positivity assumption is written as

$$0 < P(T = 1|\mathbf{X}) < 1 \tag{1.6}$$

For this assumption, all subjects have an equal chance of being exposed (or assigned to treatment). Each subject must have a treatment probability that is strictly between 0 and 1. When the probability of being treated is not possible (is equal to zero), the positivity assumption is violated.

Assumptions (1.4.1) and (1.4.2) collectively are known as the strongly ignorability as-

sumption (Rosenbaum & Rubin, 1984). Under the strongly ignorability assumption we can show that the unbiased average causal effect (β) can be estimated from observed data as follows;

$$\begin{aligned}
\beta &= E\{E[Y_1 - Y_0|X]\} \\
&= E\{E[Y_1|X] - E[Y_0|X]\} \\
&= E\{E[Y_1|T, X] - E[Y_0|T, X]\} \\
&= E\{E[Y|T = 1, X] - E[Y|T = 0, X] \qquad (1.7)
\end{aligned}$$

The proof provided in equation (1.7) takes into consideration that one has to condition on the X's. Therefore, it is essential to identify a set of confounders (**X**) that satisfy the overlap and unconfoundedness assumption.

1.5 Selection of Covariates and control for confounders

To identify the average treatment effect using the observed data, under the strongly ignorability assumption, we need to ensure that all potential confounders are identified and controlled for. There is a need to identify confounders, $\mathbf{X}$, that uphold the unconfoundedness assumption. It is possible to have covariates that may not uphold this assumption if among the identified covariates some are removed or new covariates are added, causing redundancy and hence, spurious conclusions (Austin, 2011; Austin & Stuart, 2017). Under the strongly ignorable assumption, controlling for covariates that affect both the exposure and outcome may enable us to assess the average causal effect from the observed data.

Confounding obscures the actual causal effect and results in spurious conclusions. There-

fore, it is essential to control for confounding during the design stage (such as randomization) or at the analysis stage. For observational studies, the common methods for controlling confounding are stratification and matching. In the next section, we provide a brief description of matching and stratification to reduce confounding bias.

1.5.1 Stratification

The method of stratification is commonly done during analysis. It involves creating strata according to the levels of the potential confounders. For example, in assessing the effect of appropriate complementary foods on child growth measurements such as stunting, wealth may be a potential confounder. Wealth influences the provision of appropriate foods to the child and is also associated with stunting. Therefore, we can create strata across the wealth levels, and the effect of complementary feeding on stunting is calculated in each subset. The overall average causal effect (either a mean difference, risk ratio or odds ratio) is obtained by averaging the effects across the strata. The most common approach in epidemiology and public health has been using the Mantel-Haenszel (Jager, Zoccali, Macleod, & Dekker, 2008; Kahlert, Gribsholt, Gammelager, Dekkers, & Luta, 2017). Stratification does not go without difficulty. For example, if we had ten confounders that are measured on a binary scale, then there would be $2^{10} = 1,024$ strata, and not all strata would have enough subjects ($n > 5$), as some strata would few subjects. If some strata on the exposure have very few subjects, then complete balance would not be achieved hence resulting in residual confounding which defies the very essence of stratification. Estimation of the treatment-exposure effect would be biased due to presence of some confounding and the power of the effect would be affected due to the small sample sizes.

1.5.2 Matching

Matching involves pairing exposed subjects to those unexposed on the potential confounders. For example, complementary fed and wealthy children are paired with other wealthy children, not complementary fed. Matching helps reduce the potential confounding effect of the confounders on the outcome. The problem with matching emanates when there are many confounders and the need to match on all these confounders, and the resulting sample would be very small from the original sample. Not all the exposed would be matched to the unexposed, hence, discarding those that are not matched in turn losing some information. Other methods of controlling for confounding include regression methods which use a parametric model for the conditional expectation of Y given T and X. The coefficient in the regression model correspond to the exposure effect is equal to $\beta = E[Y_1 - Y_0]$ under unconfoundedness (Senn, 1989). In most cases, linear and logistic regression are the parametric models for continuous and binary outcomes. The dimensionality problem encountered with matching and stratification is addressed using propensity score techniques Rosenbaum and Rubin (1983).

1.5.3 Instrumental variable approach

Most of the balancing methods, such as the propensity score approaches, used to control for confounding are done under the assumption of no unmeasured confounders. However, there is a possibility of having residual confounding due to unmeasured confounders. For example, in estimating the effect of child feeding intervention on child growth, we can control for factors such as age, wealth, sex, however, genetic and environmental factors are

commonly not measured. One approach that takes into account controlling for unobserved confounders is the instrumental variable approach. The instrumental variable is used to estimate causal treatment effects in the presence of measured and unmeasured covariates. An instrument is a variable that predicts an exposure, but conditional on the exposure shows no independent association with the outcome. Instrumental variables have been widely used and are categorised as either "*exogenous*" and "*endogenous*" (Sainani, 2018). An endogenous variable refers a variable that is related to the outcome variable or variables related to the outcome. On the other hand, an "exogenous" variable is a variable that is not related to any factor of interest. We illustrate the relationship of an instrumental variable with the treatment and outcome of interest in Figure 1.1.

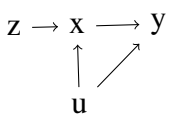


Figure 1.1: Illustration of Instrumental variable. **Source**: Sainani (2018)

In figure 1.1, u represents the unmeasured confounders and is associated with x, (for example, children living in drought prone environments are more likely to have poor child growth outcomes unlike those living in areas where the climate is favourable) hence, producing bias estimates of x on y. We identify an instrument z which is related to x but not to u, hence z is exogenous. Therefore, with the identification of the instrument z, we can obtain unbiased estimates of the effect of x on y

1.6 Data Description

1.6.1 Malawi Demographic and Health Survey (MDHS)

In this thesis, we used both the 2010 and 2015-16 Malawi Demographic and Health Survey (DHS) and 2018 Zambia DHS datasets to estimate the effect of maternal HIV on perinatal mortality, childbirth weight, and the effect of appropriate complementary feeding on child growth outcomes. These are nationally representative household surveys that provide data for a wide range of population, health, and nutrition indicators focused on under five-year-old children in Malawi and Zambia. The datasets provided up-to-date information on current HIV trends, child health, and nutrition. In addition, the 2010 data provided information regarding the moderating effect of ART status on the association between maternal HIV infection and birth weight. Parents of infants included in the surveys had signed informed consent. The MDHS and ZDHS reports provide the details of the sampling design (MDHS, 2010; Government of Malawi & ICF, 2017; Zambia Statistics Agency & ICF, 2019; Twabi, Manda, & Small, 2020).

We used three anthropometric indices for the child growth outcomes, namely height-for-age, weight-for-height, and weight-for-age. We followed the convention of recording each index as a *z-score*, which indicates how many standard deviations (SD) is a child's index from the median of the World Health Organization (WHO) Child Growth Standards (WHO, 2006). Height-for-age *z-score* $<$ -2 SD denotes a child who is stunted; similarly for wasting (weight-for-height *z-score* $<$ -2 SD), and underweight (weight-for-age *z-score* $<$ -2 SD). In measuring the appropriate complementary feeding practice exposure, we considered a child to be appropriately complementary fed if the child satisfied the following three

conditions, the child was introduced to complementary foods, the child had a minimum dietary diversity, and a minimum meal frequency (Twabi et al., 2021).

Covariates that were identified within the Demographic and Health Survey datasets as potential confounders for the effect of maternal HIV on perinatal mortality and childbirth weight included: mother's education, a child's age, place of residence, wealth. For the effect of appropriate complementary feeding on the child growth measurements, the confounders were mother's education, counselling on breastfeeding, history of diarrhoea, household wealth, child's age, mother's employment, sex of a child, place of residence, wealth quantile, mother's age, mother's HIV status, birth weight, and antenatal visits. However, we replaced mothers' employment status with mothers' education and wealth as only a few mothers were employed full-time since the matched sample would have been small within mothers belonging in the full-time category.

1.6.2 Malawi Longitudinal Study of Families and Health (MLSFH)

The Malawi Longitudinal Study of Families and Health (MLSFH) is an ongoing panel survey that has been conducted bi-annually since 1998 in 120 villages in rural areas of three districts; Rumphi, Mchinji, and Balaka in Malawi (H.-P. Kohler et al., 2014; Baranov & Kohler, 2018). The aim of the MLSFH is to understand social economic influences on fertility behaviors and HIV risk perceptions. In 1998, the MLSFH randomly selected households from which to interview ever-married women and their husbands in the three districts. Major MLSFH data collection was conducted in the years 1998, 2001, 2004, 2006, 2008, and 2010. Extensions on dynamics in the aging population was done from 2012 to 2020. MLSFH has collected socio-economic, fertility, and HIV data for up to

4000 couples in the 120 villages across the three districts. The design and methods of the MLSFH have been described in detail elsewhere (H.-P. Kohler et al., 2014; Baranov & Kohler, 2018; I. Kohler et al., 2020). We studied 2540 individuals enrolled in 2004 and followed up at eight time points namely; 2006, 2008, 2010, 2012, 2013, 2017, 2018, and 2019.

1.7 Research Problem

In biomedical and health research studies, often multiple outcomes are collected, and the interest is to estimate treatment effects on the joint outcomes. Researchers are faced with a more complex task in estimating the effects when the outcomes are collected repeatedly over time from observational data. For example, in child growth studies, under-five (U5) year old children, who are assumed to be exclusively breastfed (children less than six months of age) and complementary fed (children aged 6 to 23 months), attend U5 clinics every month where their anthropometric measurements are collected. The interest would be to assess changes in growth over time. It would therefore be essential to control for confounders when intending to estimate the causal effect of appropriate complementary feeding on child growth from the observational data. In addition, since the outcomes are measured within the same child, they are possibly correlated. Hence, the causal analysis should consider the correlation within and between the outcomes.

Several studies have developed methods for estimating exposure effects for multivariate outcomes. However, certain methods utilized standard regression adjustment to handle confounding, and some were developed specifically for randomized settings. Sammel, Lin,

and Ryan (1999) proposed a multivariate linear mixed-effects model for several outcomes, linking the separate linear mixed models using correlated model-specific random effects. However, their framework did not establish a clear means of estimating causal treatment effects. Lupparelli and Mattei (2017) employed the potential outcome framework and developed a product set for binary non-independent outcomes to evaluate the causal effect of an exposure on the joint multivariate structure of the binary outcomes. Yet, their approach was intended for a single time-point randomized setting and relied on log-mean linear models for treatment effect estimation. Mattei, Li, Mealli, et al. (2013) used a Bayesian approach to estimate the causal effect of an exposure on the main outcome, regarding the remaining outcomes as auxiliary outcomes, and hence did not specify the full correlation structure for the main and auxiliary outcomes. Hernan, Brumback, and Robins (2002) estimated the effect of time-varying exposures on observational repeated measurements with time-varying confounders using marginal structural models. However, their approach only accounted for the serial correlation between the repeated measurements. Furthermore, several methods have been developed for multiple treatment effects on an outcome for longitudinal data (Howe, Cole, Mehta, & Kirk, 2012; Tager, Haight, Sternfeld, Yu, & van Der Laan, 2004; Taubman, Robins, Mittleman, & Hernán, 2009). We searched PubMed using the keywords ”Causal inference” OR ”causal effects” AND Longitudinal data AND Marginal structural Models AND Multiple outcomes. Little is known of studies that have estimate treatment effects for longitudinal observational data with multiple outcomes while controlling for time-varying confounders and treatments and accounting for the within (serial) and between outcome correlation. VanderWeele, Mathur, and Chen (2020) provided a comprehensive description and example of causal effect estimation for longitudinal data with multiple outcomes using standard causal inference methods. Moodie, Stephens, and

Klein (2014) developed a semi-parametric survival technique using the marginal structural model for multiple-outcome survival data to assess the effect of exposure on several cause-specific deaths. Their interest was to estimate the effect of an exposure on several causes-specific deaths. Agogo, Murphy, McAvay, and Allore (2019) employed a shared random effects model to jointly analyze concurrent binary outcomes in a longitudinal observational setting but did not control for time-varying confounders.

Unlike standard longitudinal studies with a univariate outcome, the estimation of causal effects of longitudinal studies with multiple outcomes can be challenging as one needs to account for both the within and between outcome correlation as well as control for time-varying treatments and confounders, including the outcome-specific confounders. Ignoring the complex dependency structure within the estimation may inflate the standard errors and lead to overestimation of the causal effect (Oliveira & Teixeira-Pinto, 2015; Cho, 2016).

This Ph.D. study is important because it develops statistical methods to estimate treatment effects for longitudinal observational data with multiple outcomes. The methods provide an advanced approach for public health researchers to analyze non-randomized multiple outcome data collected repeatedly over time.

1.8 Research Aim

The overall aim of this Ph.D. thesis is to develop causal inference methods that could be used to estimate causal effects in longitudinal observational studies with multiple outcomes. The proposed methods are built around the Marginal Structural Model with inverse prob-

ability treatment weights (MSM-IPTWs). The proposed methods control for time-varying treatment and confounders and account for the within-and between outcome correlation. Furthermore, as part of a comprehensive simulation study, the comparative performance of statistical regression approaches and an application to maternal and child health studies from a cross-sectional survey and on the estimation of the effect of HIV status awareness on risky sexual behavior from a longitudinal MLSFH survey is done.

1.9 Outline of the Thesis

The rest of this thesis is organized as follows; Chapter 2 reviews propensity score methods based on matching, inverse weighting, and subclassification. We discuss how propensity score methods balance covariates among exposure groups and a derivation of the estimators from the observed data. We apply the PS methods to the Malawi Demographic and Health survey data. Chapter 3 presents the Marginal structural model and inverse probability treatment weight for univariate longitudinal outcomes and its application on the Malawi Longitudinal Study for Family and Health (MLSFH) data to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners. Chapter 4 presents an overview of statistical regression methods used to estimate treatment effects for multiple outcomes. We compare the performance of four multivariate methods using a simulation-based study. An empirical application is conducted to assess the effect of appropriate complementary feeding on three child growth outcomes. Chapter 5 presents our proposed methods of estimating causal inference for longitudinal observational data with multiple outcomes. We propose an estimator for balancing exposure groups based on the inverse propensity weights and modify the marginal structural model to estimate the causal effects

for the multiple outcomes. We demonstrate the applicability of the proposed methods on an empirical application to the MLSFH to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners. Finally, in Chapter 6, the contributions of the approaches are discussed and summarised. The limitations and future research are discussed.

CHAPTER 2 : Theory of Propensity Scores and Application to Child Health

2.1 Introduction

The previous chapter provided a brief overview of the propensity score and how it is used to control for confounding. This chapter provides details of the propensity score matching, inverse propensity score weighting, and propensity score subclassification, the corresponding estimators, and a derivation of their variances. We apply maternal and child health data methods from the Malawi and Zambia Demographic and Health Surveys (MDHS). The publications:Twabi, H.S., Manda, S.O. & Small, D.S. Assessing the effects of maternal HIV infection on pregnancy outcomes using cross-sectional data in Malawi. BMC Public Health 20, 974 (2020). https://doi.org/10.1186/s12889-020-09046-0 and Twabi HS, Manda SOM, Small DS. Evaluating the Effect of Appropriate Complementary Feeding Practices on Child Growth in Malawi Using Cross-Sectional Data: An Application of Propensity Score Matching. Front Nutr. 2021 Nov 18;8:714232. doi: 10.3389/fnut.2021.714232 and Twabi, H. S., Manda, S. O. M., & Small, D. S. (2022). Estimating Unhealthy Food Effects on Childhood Overweight in Malawi Using an Observational Study. Maternal and Child Health Journal, 1-9., are based on material from this chapter.

2.2 Propensity score Method

The propensity score is defined as the probability that subject i is given treatment or is exposed given observed covariates $\mathbf{X}$. The propensity score is defined mathematically as;

$$p(x) = Pr(T = 1|\mathbf{X}) \tag{2.1}$$

$$0 < p(x) < 1$$

The propensity score reduces the dimensionality of a vector of covariates $\mathbf{X}$ by providing a single score. Equation 2.1 implies that T and X are independent conditional on the propensity score $X \amalg T|p(x)$ (Rosenbaum & Rubin, 1984). This ensures that exposed and unexposed subjects with the same propensity score are balanced across the distribution of their respective confounders $\mathbf{X}$. Dimensionality reduction using the propensity score can be done using matching on the propensity score, inverse propensity score weighting (Austin & Stuart, 2015), augmented inverse probability weighting that incorporates a model for the outcome (Lunceford & Davidian, 2004), doubly robust estimators (Cao, Tsiatis, & Davidian, 2009), stratification (Rosenbaum & Rubin, 1984) and regression adjustment (Austin, 2011).

Rosenbaum and Rubin (1983) proved that the propensity score is a balancing score and conditioning on the PS, the potential outcomes are independent of the exposure for $p\varepsilon(0, 1)$.

The balancing property of the propensity score shows that $P(T = 1|p(x), x) = p(x)$

$$\begin{aligned} P(T = 1|p(x)) &= E(T|p(x)) \\ &= E[E(T|p(x), x)|p(x)] \\ &= E[E(T|x)|p(x)] \\ &= E[Pr(T = 1|x)|p(x)] \\ &= E[p(x)|p(x)] \\ &= p(x) \end{aligned}$$

The equation holds using the unconfounded assumption, iterated expectations, definition of the PS. Hence we can further define the assumption 1.4.1 provided in Chapter 1 as (Y_1, Y_0) II $T|p(x)$ and show that conditional on the propensity score, the mean potential outcomes may be identified from observed data.

$$E\{E[Y|T = 1, p(x)]\} = E\{E[Y_1|T = 1, p(x)]\} = E\{E(Y_1|p(x))\} = E(Y_1)$$

The same applies for $T = 0$, resulting in $E\{E[Y|T = 1, p(x)]\} - E\{E[Y|T = 0, p(x)]\} = E[Y_1 - Y_0]$.

2.2.1 Propensity Score Estimation

The common approach used to estimate the propensity score is assuming a logistic distribution for the PS and estimating the probability using a logistic regression model. The estimated propensity score is the predicted probability of the exposure from the fitted

regression model (Austin, 2011).

$$\widehat{p}(x)_i = \text{logit}(p(x_i)) = \log\left(\frac{p(x_i)}{1-p(x_i)}\right) = \log\left(\frac{\Pr(A_i=1|x_i)}{1-\Pr(A_i=1|x_i)}\right) = \mathbf{X}_i\boldsymbol{\beta}$$

In multivariate form we can write it as $p(\mathbf{X}, \boldsymbol{\beta}) = \{1 + \exp(-\mathbf{X}^T\boldsymbol{\beta})\}^{-1}$, where $\boldsymbol{\beta}$ is a $(p \times 1)$ vector. Interactions and higher-order terms can be included in the model. The parameter $\boldsymbol{\beta}$ is estimated by the maximum likelihood estimator (MLE).

$$\log L(T_i, \mathbf{X}_i; \boldsymbol{\beta}) = \sum_{i=1}^{n} \frac{T_i - p(\mathbf{X}_i, \boldsymbol{\beta})}{p(\mathbf{X}_i, \boldsymbol{\beta})\{1 - p(\mathbf{X}_i, \boldsymbol{\beta})\}} \frac{\partial}{\partial \beta}\{p(\mathbf{X}_i, \boldsymbol{\beta})\} = \mathbf{0} \tag{2.2}$$

An extension of the logistic regression model for the case of binary treatment would be the multinomial logistic regression model for multilevel treatments (Imbens, 2000; Feng, Zhou, Zou, Fan, & Li, 2012). For treatments or treatment levels that are correlated, Imbens (2000) recommended nested logistic models; Lechner (2002) and Imai and Dyk (2004) suggested that multinomial probit models can be used. Recently, machine learning methods have also been considered for propensity score estimation. These methods include classification and regression trees (CART), bagged CART, random forests (RF), and generalized boosted models (GBM) (McCaffrey et al., 2004), which are less biased and have better 95% confidence interval coverage than the logistic regression model for the binary case (Lee, Lessler, & Stuart, 2010). A model-averaged approach called data-adaptive matching score (DAMS) that balances propensity score estimates from parametric (logistic regression) and non-parametric (random forest) models were also suggested by Zhu, Ghosh, Mitra, and Mukherjee (2014). McCaffrey et al. (2013) recently generalized the GBM to a multilevel treatment case.

Under the strongly ignorability assumption, the propensity score estimates the unbiased average treatment effects (Rosenbaum & Rubin, 1983). Methods of propensity score matching, stratification, weighting, and covariate adjustment have been developed to facilitate the causal inference using propensity scores (Austin, 2011; Cole & Hernan, 2008; d'Agostino, 1998; Rosenbaum & Rubin, 1984; Lunceford & Davidian, 2004).

2.3 Propensity Score Matching (PSM)

One of the common methods of using propensity scores is matching. The propensity score matching involves matching exposed, and unexposed subjects based on the estimated propensity score resulting in two comparable groups, and hence, the average treatment effect can be estimated from this matched sample (Rosenbaum & Rubin, 1983).

Various matching methods have been proposed in the literature to ensure optimal matching, since using the estimated PS alone to match would be limited as the probability of observing two subjects with the same value of the propensity score is zero (Becker & Ichino, 2002). These methods are the Nearest neighbor, Caliper matching, Kernel matching, and Mahalanobis metric matching.

2.3.1 Nearest Neighbour matching

The nearest neighbor matching is based on the greedy matching algorithm that matches each subject i in the treated group with a subject $i^{'}$ in the control group by the smallest

absolute distance between their propensity scores;

$$d_i = min_j|p(\mathbf{X}_i) - p(\mathbf{X}_{i'})| \tag{2.3}$$

2.3.2 Caliper matching

Caliper matching pairs each subject i in the treated group with subject $i^{'}$ in the control group within a pre-specified caliper region b; $d_{ii'} = min_{i'}\{|p(\mathbf{X}_i) - p(\mathbf{X}_i^{'})\}| < b$. This helps reduce the risk of poor matches when the distance of the propensity scores between the matched pairs is great.

2.3.3 Mahalanobis metric matching

Mahalanobis metric matching with the propensity score matches each subject i in the treated group with a subject $i^{'}$ in the control group according to the closest mahalanobis distance calculated on the similarities of the variables:

$$d_i = min_{i'}|D_{ii'}| \tag{2.4}$$

where $D_{ii'} = (\mathbf{W}_i - \mathbf{W}_{i'})\mathbf{S}^{-1}(\mathbf{W}_i - \mathbf{W}_{i'})^T$ (Rosenbaum & Rubin, 1985). Where $\mathbf{W}$ is a combined matrix of $\{\mathbf{X}, p(\mathbf{X})\}$ and $\mathbf{S}$ is the sample variance-covariance matrix of $\mathbf{X}$ for the control group.

Several improvements have been done to the various matching methods. Dehejia and Wahba (2002) introduced radius matching, which is a form of caliper matching except that the matching is one-to-many with each subject i in the treated group being matched

with multiple subjects in the control group within a pre-specified caliper region. Pan and Bai (2015) extended the caliper matching to interval matching that matches subjects based on confidence intervals in propensity scores. Further extensions include the Mahalanobis caliper matching (Guo, Barth, & Gibbons, 2006) and the genetic matching, which are forms of the Mahalanobis metric matching that uses calipers and weighted Mahalanobis, respectively.

2.3.4 PS matching estimator

Several methods have been proposed in the literature to estimate the treatment effects from the matched sample. The simplest method to estimate the treatment effect from the matched pairs considers the mean of the pair differences written as:

$$\widehat{\delta}_{match_H} = \frac{1}{n_1} \sum_{l=1}^{n_1} Y_l^1 - \bar{Y}_l^0 \tag{2.5}$$

where l represents the n_1 matched sets and $\bar{Y}_l^0$ is the average of Q unexposed subjects in the matched set l (Austin, 2014).

Matching with replacement aims to use all the observations in the data by matching the exposed or unexposed observations of the opposite treatment assignment. To define the estimator, we start by defining the potential outcomes for the matched sets as:

$$\widehat{Y}_i^1 = \begin{cases} \frac{1}{H} \sum_{l \epsilon L_H(i,\phi)} Y_i & \text{if} \quad T_i = 0 \\ Y_i & \text{if} \quad T_i = 1 \end{cases}$$

$$\widehat{Y}_i^0 = \begin{cases} Y_i & \text{if} \quad T_i = 0 \\ \frac{1}{H}\sum_{l \epsilon L_H(i,\phi)} Y_i & \text{if} \quad T_i = 1 \end{cases}$$

$L_H(i,\phi)$ is the set of H observations in the treated group opposite to i with similar propensity scores. Using these definitions of the potential outcomes, the matching replacement estimator can be written as;

$$\widehat{\delta}_{matchrep} = \frac{1}{n}\sum_{i=1}^{n}(\widehat{Y}_i^1 - \widehat{Y}_i^0)$$
$$= \frac{1}{n}\sum_{i=1}^{n}(2T_i - 1)\left(1 + \frac{Q_H(i)}{H}\right)Y_i \tag{2.6}$$

where $Q_H(i)$ is the number of times a subject i is used as a match (Abadie & Imbens, 2006). For the ATT, the estimator can be estimated to be

$$\widehat{\delta}_{matchrep(ATT)} = \frac{1}{n_1}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i^0)$$
$$= \frac{1}{n_1}\sum_{i=1}^{n} T_i - (1 - T_i)Q_H(i)Y_i \tag{2.7}$$

Abadie and Imbens (2006) proposed an estimator of the variance for the matching with replacement estimator given as;

$$\widehat{Var}(\widehat{\delta}_{matchrep}) = \frac{1}{n^2}\sum_{i=1}^{n}(\widehat{Y}_i^1 - \widehat{Y}_i^0 - \widehat{\delta}_{matchrep})^2$$
$$+\frac{1}{n^2}\sum_{i=1}^{n}\left[\left(\frac{Q_H(i)}{H}\right) + \left(\frac{2H-1}{H}\right)\left(\frac{Q_H(i)}{H}\right)\right]\widehat{\sigma}^2(X_i, T_i) \tag{2.8}$$

Where H is the number of matches per subject, $Q_H(i)$ is the number of times subject i is used as a match, and $\widehat{\sigma}^2(X_i, T_i)$ is the estimated conditional variance given by (Lunceford

& Davidian, 2004);

$$\widehat{\sigma}^2(X_i, T_i) = \frac{J}{J+1}\left(Y_i - \frac{1}{J}\sum_{l=1}^{L} Y_{s_l(i)}\right)^2 \tag{2.9}$$

where L is the fixed number of similar subjects and $s_l(i)$ is the lth closest subject to subject i among subjects with the same T value (Abadie & Imbens, 2006).

2.4 Inverse probability weighting (IPW) on the PS

Another balancing approach is weighting each subject based on the propensity score. The approach re-weight the exposed and unexposed subjects to make them representative of the population of interest. The PS weighting is similar to the approach for survey weighting (Horvitz & Thompson, 1952). The PS weighting method creates a pseudo-population in which the distribution of the confounders are balanced across the treatment groups. A treated subject is assigned a weight of $w_i = \frac{1}{p(x_i)}$, and an untreated subject is given a weight of $w_i = \frac{1}{1-p(x_i)}$. The weighted sample can then be used to estimate the exposure effect on the outcome. The weighted sample represents subjects of which their exposure is unconfounded. The inverse probability weights can be defined with respect to the exposure variable as follows

$$w_i = \frac{T_i}{p(x_i)} + \frac{(1-T_i)}{1-p(x_i)}, T_i = 0, 1 \tag{2.10}$$

Therefore the mean potential outcomes for the treated and control groups can be estimated using 2.10 as $E\{\frac{TY}{p(x)}\} = E\{\frac{TY_1}{p(x)}\}$. Under the consistency assumption, Y_1 is observed if $T = 1$ and Y_0 is observed if $T = 0$. Therefore, under the strongly ignorability assumption,

the potential outcomes can be estimated from the observed data by;

$$
\begin{aligned}
E\left\{\frac{TY}{p(x)}\right\} &= E\left\{E\left[\frac{I(T=1)Y_1}{p(x)}\middle| Y_1, \mathbf{X}\right]\right\} \\
&= E\left\{\frac{Y_1}{p(x)}E[I(T=1)|Y_1, \mathbf{X}]\right\} \\
&= E\left\{\frac{Y_1}{p(x)}P[T=1|Y_1, \mathbf{X}]\right\} \\
&= E\left\{\frac{Y_1}{p(x)}P[T=1|\mathbf{X}]\right\} \\
&= E\left\{\frac{Y_1}{p(x)} * p(x)\right\} \\
&= E[Y_1]
\end{aligned}
$$

Denoting the indicator of the treatment $T = I(T = 1)$. The same applies to $E\{\frac{TY}{p(x)}\} = E\{\frac{TY_0}{p(x)}\} = E[Y_0]$. Therefore, the average treatment effect using IPW defined by Rosenbaum and Rubin (1985) is denoted as;

$$
\begin{aligned}
\widehat{\theta}_{ipw1} &= \mu_{1,ipw} - \mu_{0,ipw} \\
&= E\left[\frac{TY}{p(x)} - \frac{(1-T)Y}{p(x)}\right] \\
&= E[\frac{TY}{p(x)}] - E[\frac{(1-T)Y}{p(x)}] \\
&= n^{-1}\sum_{i=1}^{n}\frac{T_iY_i}{\widehat{p(x)}} - n^{-1}\sum_{i=1}^{n}\frac{(1-T_i)Y_i}{1-\widehat{p(x)}}
\end{aligned}
\tag{2.11}
$$

The weights in 2.11 do not add to one, hence (Lunceford & Davidian, 2004) proposed a normalised estimator for the average treatment effect where the weights add to one for each

group as;

$$
\begin{aligned}
E\left[\frac{T}{p(\mathbf{X})}\right] &= E\left(E\left(\frac{T}{p(x)}|\mathbf{X}\right)\right) \\
&= E\left(\frac{1}{p(x)}E[T|\mathbf{X}]\right) \\
&= E\left(\frac{1}{p(x)}P(T|X)\right) \\
&= E\left(\frac{1}{p(x)}P(x)\right) \\
&= E(1) \\
&= 1
\end{aligned} \tag{2.12}
$$

Equation 2.12 applies for the treated subjects. The same proof given in Equation 2.12 applies for those who untreated $E\left[\frac{(1-T)}{(1-p(\mathbf{X}))}\right] = 1$. Thus, the second version of the IPW estimator is defined as;

$$
\widehat{\theta}_{ipw2} = \left(\sum_{i=1}^{n}\frac{T_i}{\widehat{p}(\mathbf{X})}\right)^{-1}\sum_{i=1}^{n}\frac{T_iY_i}{\widehat{p}(\mathbf{X})} - \left(\sum_{i=1}^{n}\frac{(1-T_i)}{(1-\widehat{p}(\mathbf{X}))}\right)^{-1}\sum_{i=1}^{n}\frac{(1-T_i)Y_i}{1-\widehat{p}(\mathbf{X})} \tag{2.13}
$$

$\widehat{p}(\mathbf{X})$ is estimated using a logistic regression model. The estimators 2.11 and 2.13 are solutions to the following equations (Lunceford & Davidian, 2004), we denote $p(\mathbf{X})$ as p_i;

$$
\sum_{i=1}^{n}\left\{\frac{T_i(Y_i-\mu_1)}{p_i}\right\} + \eta_1\left(\frac{T_i-p_i}{p_i}\right) = 0, \tag{2.14}
$$

and

$$
\sum_{i=1}^{n}\left\{\frac{(1-T_i)(Y_i-\mu_0)}{(1-p_i)}\right\} + \eta_0\left(\frac{T_i-p_i}{(1-p_i)}\right) = 0 \tag{2.15}
$$

Equation 2.14 and 2.15 are obtained under the assumption that the propensity score is known and letting $(\eta_0, \eta_1) = (\mu_0, \mu_1)$ to obtain $\widehat{\theta}_{ipw1}$, while $(\eta_0, \eta_1) = (0, 0)$ to yield $\widehat{\theta}_{ipw2}$.

Robins, Rotnitzky, and Zhao (1994) proposed a new class of estimators based on the inverse probability weighted estimating equations that are asymptotically normal and unbiased known as the doubly robust estimator. This weighting estimator attempts to model the outcome parametrically (Lunceford & Davidian, 2004). The doubly robust variance estimator for Δ_{ATE} is written as;

$$\widehat{\theta}_{DR} = \frac{1}{n}\sum_{i=1}^{n}\frac{T_iY_i - (T_i - \widehat{p}_i)m_1(\mathbf{X}_i, \widehat{\alpha}_1)}{\widehat{p}_i} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-T_i)Y_i - (T_i - \widehat{p}_i)m_0(\mathbf{X}_i, \widehat{\alpha}_0)}{(1-\widehat{p}_i)} \quad (2.16)$$

where $m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1)$ and $m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)$ are the predicted outcomes for the treated ($E(Y|T, \mathbf{X})$) and untreated ($E(Y|(1-T), \mathbf{X})$) groups specified by regression models. The θ_{DR} requires that both the outcome and propensity score model are correctly specified to result in the smallest large sample variance of any weighting type-estimator (Lunceford & Davidian, 2004).

The variance of the IPW estimators are derived by handling the estimates as solutions to a set of estimating equations as Hernán and Robins (2020) recommended fitting generalized estimating equations using the sandwich variance estimator. Lunceford and Davidian (2004) included the propensity score in the estimating equations and is known to be exact. Using the theory of M-estimation which implies that for the IPW estimators $\widehat{\theta}$, $n^{1/2}(\widehat{\theta} - \theta)$ converges in distribution to a $N(0, \Sigma)$ variable. We can show that the large sample variance of IPW estimators assuming the PS is known is derived as;

$$\Sigma^*_{ipw1} = E\left[\frac{(Y^1)^2}{p} + \frac{(Y^0)^2}{1-p}\right] - \widehat{\theta}_0^2 \quad (2.17)$$

and

$$\Sigma^*_{ipw2} = E\left[\frac{(Y^1-\mu_1)^2}{p} + \frac{(Y^0-\mu_0)^2}{1-p}\right] - \widehat{\theta}_0^2 \tag{2.18}$$

where $\widehat{\theta}_0$ is the true value, $\mu_1 = E[Y^1]$, $\mu_0 = E[Y^0]$.

Now, assuming that we estimate the propensity scores using the logistic regression model and we estimate $\boldsymbol{\beta}$ by maximum likelihood by solving 2.2, therefore, using M-estimators, the estimator $\widehat{\theta}_{ipw1}$ will be;

$$\boldsymbol{\phi}(\mathbf{Y}_i, \widehat{\boldsymbol{\theta}}) = \begin{pmatrix} \frac{(Y^1)^2}{p} + \frac{(Y^0)^2}{1-p} \\ \frac{T_i - p}{p\{1-p\}} \end{pmatrix}$$

where $p = p(\mathbf{X}_i, \boldsymbol{\beta})$. Therefore, the large sample variances of $\widehat{\theta}_{ipw1}$ and $\widehat{\theta}_{ipw2}$ then become

$$\Sigma_{ipw1} = \Sigma^*_{ipw1} - \mathbf{H}^T_{\beta,1}\mathbf{E}^1_{\beta,\beta}\mathbf{H}_{\beta,1}, \tag{2.19}$$

where

$$\mathbf{H}_{\beta,1} = E\left[\left(\frac{Y^1}{p} + \frac{Y^0}{1-p}\right)p_\beta\right]$$

$$\Sigma_{ipw2} = \Sigma^*_{ipw2} - \mathbf{H}^T_{\beta,2}\mathbf{E}^1_{\beta,\beta}\mathbf{H}_{\beta,2} \tag{2.20}$$

where

$$\mathbf{H}_{\beta,2} = E\left[\left(\frac{Y^1-\mu_1}{p} + \frac{Y^0-\mu_0}{1-\pi}\right)p_\beta\right]$$

$$\Sigma_{DR} = \Sigma^*_{ipw2} - E\left[\sqrt{\frac{1-p}{p}}(E[Y_1|\mathbf{X}]-\mu_1) + \sqrt{\frac{p}{1-p}}(E[Y_0|\mathbf{X}]-\mu_0)\right]^2 \tag{2.21}$$

where $p_\beta = \frac{\partial}{\partial\beta}p(\mathbf{X}, \boldsymbol{\beta})$ and $E_{\beta\beta} = E\left[\frac{p_\beta \pi'_\beta}{p(1-p)}\right]$. It has been shown that estimating $\boldsymbol{\beta}$ leads to a smaller large sample variance than using the true value of $\boldsymbol{\beta}$ (Lunceford & Davidian,

2004).

The components in 2.4-2.21 can be estimated using the empirical estimators of $\mathbf{A_n}(\theta_0)$ and $\mathbf{B_n}(\theta_0)$. Using these empirical estimates yields the sandwich variance estimator. The empirical sandwich variance estimators for the IPW estimators are given as;

$$\frac{1}{n^2}\sum_{i=1}^{n}\widehat{I}_i^2, \tag{2.22}$$

where

$$\widehat{I}_{ipw1,i} = \frac{T_iY_i}{\widehat{p}_i} - \frac{(1-T_i)Y_i}{1-\widehat{p}_i} - \widehat{\theta}_{ipw1} - (T_i-\widehat{p}_i)\widehat{\mathbf{H}}'_{\beta,1}\widehat{\mathbf{E}}^{-1}_{\beta,\beta}\mathbf{X}_i$$

$$\widehat{I}_{ipw2,i} = \frac{T_i(Y_i-\widehat{\mu}_{1,ipw2})}{\widehat{p}_i} - \frac{(1-T_i)(Y_i-\widehat{\mu}_{0,ipw2})}{1-\widehat{p}_i} - (T_i-\widehat{p}_i)\widehat{\mathbf{H}}'_{\beta,2}\widehat{\mathbf{E}}^{-1}_{\beta,\beta}\mathbf{X}_i,$$

$$\widehat{I}_{DR,i} = \frac{T_iY_i - m_1(\mathbf{X}_i,\widehat{\alpha_1})(T_i-\widehat{p}_i)}{\widehat{p}_i} - \frac{(1-T_i)Y_i + m_0(\mathbf{X}_i,\widehat{\alpha_0})(T_i-\widehat{p}_i)}{1-\widehat{p}_i} - \theta_{DR},$$

where

$$\mathbf{E}^{-1}_{\beta,\beta} = \frac{1}{n}\sum_{i=1}^{n}\widehat{p}_i(1-\widehat{p}_i)\mathbf{X}_i\mathbf{X}_i',$$

$$\widehat{\mathbf{H}}_{\beta,1} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{T_iY_i(1-\widehat{p}_{ii})}{\widehat{p}_i} - \frac{(1-T_i)Y_i\widehat{p}_i}{1-\widehat{p}_i}\right\}\mathbf{X}_i,$$

$$\widehat{\mathbf{H}}_{\beta,2} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{T_i(Y_i-\widehat{\mu}_{1,ipw2})(1-\widehat{p}_{ii})}{\widehat{p}_i} - \frac{(1-T_i)(Y_i-\widehat{\mu}_{0,ipw2})\widehat{p}_i}{1-\widehat{p}_i}\right\}\mathbf{X}_i,$$

2.5 Propensity Score Sub-classification

Stratification on the propensity score, also known as sub-classification on the PS, involves ranking subjects according to their estimated PS (Cochran, 1968; Austin, 2011). One can assume that each stratum mimics a randomized situation within which the distributions

of the measured covariates are balanced in expectation. The number of strata that can be used can vary. However, five strata are recommended because they can reduce 90% of bias (Rosenbaum & Rubin, 1984; Lunceford & Davidian, 2004). Supposed we divide a sample into p groups, where the q^{th} group consists of subjects with propensity scores falling within a subset $I_q = (d_{q-1}, d_q]$. If the assumption of strongly ignorable treatment assumption holds, unbiased average treatment effects for $T = 0, 1$ can be estimated within each stratum (Rosenbaum & Rubin, 1984),

$$\widehat{E}(y_{i,1}|d_i \epsilon I_q) = \frac{\Sigma_{i:d_i \epsilon I_q, T_i=1} y_{i,1}}{n_{q1}}, \widehat{E}(y_{i,0}|d_i \epsilon I_q) = \frac{\Sigma_{i:d_i \epsilon I_q, T_i=0} y_{i,0}}{n_{q0}} \tag{2.23}$$

Where n_{q1} and n_{q0} are the number of subjects in the q^{th} subgroup for the treatment and control group respectively. So the treatment effect for the q^{th} subgroup an be estimated by

$$\widehat{E}(y_{i,1}|d_i \epsilon I_q) - \widehat{E}(y_{i,0}|d_i \epsilon I_q) \tag{2.24}$$

The overall treatment effect can then be estimated by pooling the estimated effects across the strata. A common approach is to use the Mantel-Haenszel to pool the effects across the strata (Kahlert et al., 2017), which can be done using epidemiological packages in available software. Within each stratum, treated and untreated subjects will have similar values of the PS. Therefore when the PS has been correctly specified, the distribution of measured baseline covariate will be approximately identical between treated and untreated subjects within the same stratum. Studies have shown using Monte Carlo simulation and empirical examples that matching and inverse probability weighting eliminate systematic differences between the treated and untreated subjects to a greater degree than stratification or covariate adjustment. Furthermore, the estimates of the effects after applying PS

matching and weighting were close to those of randomized control trials (RCTs) (Kurth et al., 2006).

Most of the propensity score methods have been applied to binary treatment. Imbens extends Rosenbaum and Rubin's work to multiple treatments (Imbens, 2000). He defines the generalized propensity score as the conditional probability of receiving a particular level of the treatment given the pre-treatment variables (baseline covariates). Extensions of the generalized propensity score include matching, weighting, and subclassification (Gao, Zhang, Liang, Sun, & Wang, 2021; Austin, 2018). The following section illustrates the propensity score matching, weighting, and subclassification on the Malawi Demographic and Health Survey (MDHS) and the Zambia Demographic and Health Survey (ZDHS) survey data.

2.6 Application of PS matching and Weighting on Child Health Data

2.6.1 Example 1: Assessing the effect of Maternal HIV on pregnancy outcomes using cross-sectional data in Malawi

Maternal HIV infection is associated with increased rates of adverse pregnancy outcomes such as low birth weight (LBW) (weighing less than 2500g at birth) and increased perinatal, and early neonatal mortality (Dreyfuss et al., 2001; Rollins et al., 2007). Ascertaining the causal association between maternal HIV infection and birth weight and perinatal mortality could be conflicted by the association between ART and the outcomes in HIV-infected mothers. Therefore, the study aimed to assess the effect of maternal HIV infection on birth weight and perinatal mortality.

Data were extracted from the 2010 and 2015-16 Malawi Demographic and Health Surveys (MDHS). The study analyzed data on 4,111 and 4,759 children born to mothers who had HIV during the respective surveys. The outcome variables of interest were: Birth weight (grams) and perinatal mortality (0= Alive 1= Dead). The exposure variable of interest was mothers' HIV-infection status (0 = HIV negative, 1 = HIV positive). The moderator variable was ART status among HIV-infected mothers (0 = Not on ART, 1 = On ART) using the 2010 MDHS.

2.6.1.1 Results on maternal HIV effects on child birth outcomes

Table 2.1 presents results for a comparison of the effect of maternal HIV infection, birth weight, and perinatal mortality before and after applying the PS methods. Tables that provide the balance of confounders after matching are shown elsewhere (Twabi et al., 2020). The results for the 2010 data have shown that maternal HIV infection was negatively associated with birth weight. Infants of HIV-infected mothers had a lower average birth weight of -25.3g (95% CI: -95.5, -7.4) than infants of HIV-uninfected mothers.

The 2015-16 results have shown a significant positive association between maternal HIV infection status and birth weight. Infants of HIV-infected mothers had a higher average birth weight, 116.3g (95% CI: 27.8, 204.7), than infants of HIV-uninfected mothers. In addition, a comparison of the odds ratio of perinatal mortality between HIV-infected and HIV-uninfected mothers showed an increase in perinatal deaths in children born HIV-infected mothers in 2010. The odds of perinatal mortality were 1.5 times greater among infants of HIV-infected mothers than infants of HIV-uninfected mothers. However, in 2015, there was no significant difference in perinatal mortality between HIV-infected and

HIV-uninfected mothers.

Table 2.1: The effect of maternal HIV-infection on birth weight and perinatal mortality

	Child Health Outcomes 2010		**Child Health Outcomes 2015-16**	
	Birth Weight	**Perinatal Mortality**	**Birth Weight**	**Perinatal Mortality**
	Mean.Diff(95% CI)	OR(95% CI)	Mean.Diff(95% CI)	OR(95% CI)
Unadjusted	-97.2(-160.4, -34)	1.8(1.2, 2.8)	52.6 (41.6, 125.5)	1.4(0.8, 2.6)
PS matching	-149(-324.2,-26.2)	2.4(1.2,4.8)	144.7(40.4,249.1)	1.7(0.8,3.8)
IPW	-25.3(-95.5, -7.4)	1.5(1.1, 3.1)	116.3(27.8, 204.7)	1.0(0.4, 1.6)
Subclassification	34.2 (-39.4, 107.9)	1.16 (0.7, 2.0)	86.6 (35.6, 137.6)	0.97 (0.6, 1.7)

Table 2.2 presents results for the effect of maternal HIV infection on birth weight as a binary outcome. Overall for both datasets, infants born to HIV-infected mothers were more likely to have low birth weight than those born to HIV-uninfected mothers. A statistically significant effect was observed for the inverse propensity score weighting only. Using the IPW, in 2010, children born to HIV-infected mothers were more likely to have a low birth weight than children born to HIV-uninfected mothers. However, in 2015-16, children born to HIV-infected mothers were less likely to have a low birth weight as compared to children born to HIV-uninfected mothers. The IPW had narrower confidence intervals than those for the standard regression, PS matching, and subclassification methods.

Table 2.2: The effect of maternal HIV-infection on binary child birth weight

	2010 MDHS	**2015-16 MDHS**
	OR(95% CI)	**OR(95% CI)**
Adjusted covariates	1.3(0.9,1.7)	1.2(0.8,1.6)
PS matching	1.3(0.9,1.9)	1.0(0.7,1.6)
IPW	1.2(1.0,1.4)	0.9(0.8,0.95)
Subclassification	1.2 (0.9, 1.6)	1.1 (0.8, 1.5)

2.6.2 Example 2: Association between exclusive breastfeeding and appropriate complementary feeding and anthropometric measures among infants and young children in Malawi and Zambia

Child malnutrition has been shown to adversely affect child growth (WHO, 2009). The WHO recommends exclusive breastfeeding and appropriate complementary feeding as the optimal infant feeding practice for children under two years of age to ensure good child growth and development (Dewey, 2003; WHO & UNICEF., 2003; WHO, 2010). Despite the efforts to implement the policies and guiding principles on infant and young child feeding, the prevalence of under-nutrition remains high in Malawi as compared to other Sub-Saharan African (SSA) countries (Government of Malawi & ICF, 2017; WHO, 2017). Although previous studies have assessed child complementary feeding and child growth in Malawi, including in an HIV context (Vaahtera et al., 2001; Chinkonde, Hem, & Sundby, 2012; Chintalapudi et al., 2018), these studies used observational data which are prone to confounder bias, and balancing methods such as the propensity score methods were rarely used. Therefore, the study aimed to assess the causal association between appropriate complementary feeding and child growth using statistical tools that control for confounding. This study provides an understanding of how purported infant and young child feeding impacts child growth, which in turn provides reliable evidence to policy-makers

Child data from the 2015-16 Malawian Demographic Health Survey (MDHS) and the 2018 Zambian Demographic and Health Survey (ZDHS) were used. The study analyzed data on 2059 and 1231 children aged 0-6 months in Malawi and Zambia, respectively, and had exclusive breastfeeding outcomes. Furthermore, data on 4762 children aged 6 to 24 months

from the MDHS data and 2784 from the ZDHS data, which had an appropriate complementary feeding outcome, were analyzed. The exposure variable was exclusive breastfeeding (0= Exclusive breastfed, 1= Not Exclusive breastfed) and appropriate complementary feeding (CF) categorized as (0= Not appropriate, 1 =Appropriate). The study used the three most common anthropometric indices of measuring a child's growth height-for-age *z-score* (HAZ), weight-for-height *z-score* (WHZ) and weight-for-age *z-score* (WAZ) (De Onis, Blossner, Organization, et al., 1997).

2.6.2.1 Results on child feeding effects on growth

Table 2.3 presents the effects of exclusive breastfeeding and appropriate complementary feeding on the binary child growth outcomes before applying the propensity score methods on the data. Binary estimates for the 2015-16 MDHS showed that exclusively breastfed children were less likely to be stunted or underweight, and were more likely to be wasted. However, the association was significant for underweight only. Using the 2018 ZDHS, there was no significant association between exclusive breastfeeding and the growth measurements. However, children who were provided appropriate complementary foods were less likely to be underweight or wasted. In Malawi, appropriate complementary feeding was not associated with the growth measurements.

Table 2.4 present the continuous estimates of EBF and ACF. Using the 2015-16 MDHS, we observe that children who were exclusively breastfed had an increase in their weight-for-height z-scores than those who were not exclusively breastfed. On the other hand, children who were given appropriate complementary foods had increased weight-for-age z-scores compared to children who were given unhealthy foods. Using the 2018 ZDHS, an increase

in height-for-age z-scores and weight-for-age z-scores was associated with exclusively breastfeeding. Additionally, Children who were given appropriate complementary foods had increased height-for-age z-scores compared to those who had inappropriate foods.

Table 2.3: Adjusted effect of exclusive breastfeeding and appropriate complementary feeding on binary child growth indicators before matching and IPW

	2015-16 MDHS		**2018 ZDHS**	
	EBF	**Appropriate complementary**	**EBF**	**Appropriate complementary**
	OR(95% CI)	**OR(95% CI)**	**OR(95% CI)**	**OR(95% CI)**
Stunting	0.9(0.5,1.7)	1.3(0.5,3.4)	1.4(0.5,3.7)	0.8(0.6,1.2)
Underweight	0.2(0.1,0.7)	0.7(0.4,1.2)	0.9(0.5,1.6)	0.7(0.6,0.9)
Wasted	1.1(0.5,2.2)	0.6(0.3,1.1)	0.4(0.1,1.4)	0.7(0.5,0.8)

Table 2.4: Adjusted effect of exclusive breastfeeding and appropriate complementary feeding on continuous child growth indicators before matching and IPW

	2015-16 MDHS		**2018 ZDHS**	
	EBF	**Appropriate complementary**	**EBF**	**Appropriate complementary**
	Coef.(95% CI)	**Coef.(95% CI)**	**Coef.(95% CI)**	**Coef.(95% CI)**
Height-for-age z-score	-0.7(-1.7,0.3)	0.3(-0.01,0.5)	0.2(0.04, 0.4)	0.2(0.02,0.4)
Weight-for-age z-score	-0.6(-1.5,0.4)	0.24(0.01,0.47)	0.23(0.1,0.4)	0.03(-0.11,0.2)
Weight-for-height z-score	1.4(0.03,2.9)	-0.03(-0.8,0.7)	1.0(-0.03,2.1)	-0.1(-0.2,0.1)

After applying the PS methods and estimating the effects using the three PS approaches, we find that children from Zambia who were exclusively breastfed were less likely to be stunted using the PS matching and weighting, as shown in Table 2.5. The effects vary with the PS methods, with the PS weighting providing a higher effect (OR=0.6, (95% CI: 0.4,0.9)) than the PS weighting and subclassification. In addition, the estimates for the PS matching had narrower confidence intervals. However, there was no association between

exclusive breastfeeding and underweight and wasting across the PS techniques.

Using the MDHS data, children who were exclusively breastfed had reduced effect of being underweight and stunted. A statistically significant effect of EBF was observed for underweight and stunting (OR=0.3, (95% CI: 0.1, 0.7)) using the PS matching and (OR=0.2, (95% CI: 0.1, 0.5)) inverse PS weighting. The estimates after applying the PS techniques are in contrast to those obtained before applying the PS methods. We observe a significant reduction in stunting and underweight for specific PS methods, mostly PS matching and IPW. Before applying the PS techniques, estimates showed no significant effect of EBF on underweight and wasting and an increase in wasting among EBF children

For the continuous outcomes, overall for both the MDHS and ZDHS data, children who were exclusively breastfed had an increase in their height-for-age *z-scores* as compared to children who were not exclusively breastfed using PS matching. However, only for the ZDHS data did exclusive breastfeeding lead to an increase in weight-for-age z-scores. These findings may signify improved growth in children who were exclusively breastfed.

Table 2.5: Exclusive breastfeeding effect on binary child growth outcomes after applying PS approaches (MDHS and ZDHS)

	2015-16 MDHS			**2018 ZDHS**		
	Wasting	**Underweight**	**Stunting**	**Wasting**	**Underweight**	**Stunting**
	OR(95%CI)	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**
PS Matching	0.8(0.3,2.1)	0.3(0.1,0.7)	0.6(0.4,0.9)	1.1(0.4,2.3)	0.6(0.2,1.6)	0.5(0.3,0.8)
IPW	0.6(0.3,1.2)	0.2(0.1,0.5)	0.5(0.3,0.7)	1.1(0.6, 2.3)	0.7(0.3, 1.7)	0.6(0.4, 0.9)
Subclassification	0.5(0.2,1.4)	0.7(0.3,1.5)	0.9(0.5,0.1.7)	0.9(0.5, 1.7)	0.6(0.4, 1.0)	0.8(0.7, 1.2)

Table 2.6: Exclusive breastfeeding effect on continuous child growth outcomes after PS Matching and IPW for the MDHS and ZDHS data

	2015-16 MDHS			2018 ZDHS		
	WHZ	**WAZ**	**HAZ**	**WHZ**	**WAZ**	**HAZ**
	Coef.(95%CI)	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**
PS Matching	0.3(0.05,0.6)	0.2(-0.02,0.4)	0.07(-0.2,0.3)	1.3(0.1,2.4)	**0.2(0.1,0.4)**	0.1(-0.1,0.3)
IPW	0.3(-0.04,0.6)	0.2(-0.1,0.4)	-0.04(-0.4,0.3)	0.1(-1.1,1.1)	-0.1(-0.2,0.1)	0.04(-0.2,0.3)
Subclassification	0.1(-0.1,0.2)	0.05(-0.1,0.2)	-0.004(-0.2,0.2)	0.04(-0.07,0.2)	0.1(-0.04,0.2)	0.03(-0.1,0.1)

Tables 2.7 presents estimates for the effect of appropriate complementary feeding on wasting, stunting and underweight and Table 2.8 presents estimates for the effect of appropriate complementary feeding on weight-for-height, height-for-age and weight-for-age *z-scores*. The MDHS data showed that children on appropriate complementary feeding were less likely to be underweight using the inverse PS weighting. A statistically significant effect was observed for underweight only. For the ZDHS data, children who were given appropriate complementary foods were less likely to be stunted (OR=0.2, (95% CI:0.1, 0.6)) and underweight (OR=0.8, (95% CI:0.7, 0.9)), using the PS matching, compared to children who were not given appropriate complementary foods. We note that the appropriate complementary effects before applying the PS methods were either underestimated or estimated, especially for the effect of ACF on wasting using the ZDHS data. The estimates for the PS methods were slightly different, with a higher effect observed using the PS matching and subclassification compared to estimates for the IPW.

For the continuous outcomes, using the PS matching and IPW, children who were appropriately complementary fed increased the weight-for-age z-score using the MDHS data. While, for the ZDHS, across the PS methods, children who were appropriately com-

plementary fed seemed to have a decrease in their height-for-age, weight-for-age, and weight-for-height z-scores. However, for both datasets and across all the PS methods, the effects were not significant.

Table 2.7: Effect of appropriate complementary feeding on binary child growth outcomes for children aged 6-23 months for the MDHS and ZDHS data

	2015-16 MDHS			**2018 ZDHS**		
	Wasting	**Underweight**	**Stunting**	**Wasting**	**Underweight**	**Stunting**
	OR(95%CI)	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**	**OR(95%CI)**
PS Matching	0.9(0.8,1.2)	0.8(0.5,1.4)	0.9(0.6,1.4)	1.1(0.7,1.6)	0.8(0.7,0.9)	0.7(0.6,0.8)
IPW	0.7(0.2,3.0)	0.2(0.1,0.6)	0.4(0.1,2.8)	1.2(0.6, 2.3)	0.8(0.4, 1.6)	0.9(0.4, 0.95)
Subclassification	0.9(0.4,2.6)	0.7(0.4,1.4)	0.8(0.5,1.2)	0.9(0.5, 1.7)	0.6(0.4, 1.0)	0.9(0.7, 1.2)

Table 2.8: Effect of complementary feeding on continuous child growth outcomes for children aged 6-23 months for the MDHS and ZDHS data

	2015-16 MDHS			**2018 ZDHS**		
	WHZ	**WAZ**	**HAZ**	**WHZ**	**WAZ**	**HAZ**
	Coef.(95%CI)	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**	**Coef.(95%CI)**
PS Matching	-0.1(-0.4,0.2)	0.03(-0.24,0.3)	0.15(-0.22,0.53)	-0.1(-0.3,0.1)	-0.1(-0.23,0.13)	0.04(-0.2,0.3)
IPW	0.01(-0.2,0.2)	0.15(-0.1,0.4)	0.3(-0.1,0.7)	-0.1(-0.2,0.1)	-0.02(-0.2,0.2)	-0.02(-0.3,0.2)
Subclassification	0.1(-0.2,0.3)	0.17(-0.04,0.4)	0.24(-0.004,0.5)	-0.02(-0.1,0.1)	-0.02(-0.2,0.1)	-0.03(-0.2,0.1)

2.7 Summary

In this chapter, we have reviewed and discussed propensity score methods. Using Demographic and Health Surveys from Malawi and Zambia, we have illustrated the propensity score methods on maternal and child health data.

We have found that in Malawi, maternal HIV infection impacts perinatal mortality and childbirth weight, with children born to HIV mothers being more likely to have low birth weight and high perinatal mortality using the 2010 MDHS. However, we observe a significant increase in childbirth weight among children born to HIV-infected mothers

for the 2015-16 MDHS. This effect may signify the effectiveness of PMTCT intervention between 2010 and 2015-16 in Malawi. In addition, we found that for the 2015-16 MDHS and the 2018 ZDHS, exclusive breastfeeding and appropriate complementary feeding result in a reduction of adverse effects of stunting, wasting and underweight and also result in an increase in the child growth measurements.

We have observed the importance of applying the PS matching and IPW to control confounder bias when estimating the causal effect from observational studies. Before applying the PS methods, the estimates for the effect of exclusive breastfeeding and appropriate complementary feeding on the child growth outcomes are either underestimated or overestimated and show no significant effect of the exposure. In contrast, the PS methods improved the power to detect the effects. Estimating the propensity score and its balancing property is done assuming that there are no unmeasured confounders. Since we cannot assess the assumption from the data, conducting a sensitivity analysis is essential to assess how the estimated effect would change if unmeasured confounders were present. For the examples provided in this chapter, a sensitivity analysis was done, and results are described elsewhere (Twabi et al., 2020, 2021). It is important to note that all the approaches based on propensity scores can only address observed measured confounders.

Although a separate analysis of the child growth outcomes was analyzed in this chapter, since the outcomes characterize the same phenomenon (for example child growth) within a child, the outcomes are likely to be correlated. The next chapter compares multivariate methods that assess treatment effectiveness and an application using the child growth data.

CHAPTER 3 : Marginal Structural Models for repeated measures outcome data

3.1 Introduction

In the previous chapters, we assessed the causal association from cross-sectional data. However, to assess causal associations, there is a need for a temporal relationship between the exposure and the repeated outcome measurements. Longitudinal studies provide the temporal dependence between the exposure and outcome to estimate the causal effect. Longitudinal studies provide a problem of the presence of complex confounding such as those that are time-varying or the past exposure affecting future confounders (treatment-confounder feedback). Thus, there is a need to adjust for such complex confounding, at the same time accounting for the temporal dependency between the repeated outcome measurements. In the presence of time-dependent confounding and exposure history, standard balancing approaches presented in chapters 1 and 2 would not be sufficient. This chapter discusses statistical approaches for estimating treatment effects for univariate longitudinal outcomes. This chapter discusses the Marginal structural model and the inverse probability weighting estimator and applies the MSM to assess the causal effect of HIV positivity awareness on condom use and multiple sexual partners.

The paper *Twabi, H. S., Manda, S. O., Small, D. S., & Kohler, H. P. (2022). A marginal structural model for estimation of the effect of HIV positivity awareness on risky sexual behaviour. Biostatistics & Epidemiology, 6(2), 309-326.* is based on this chapter.

3.2 Notation for repeated measures data

Suppose subject i, $i = 1, \ldots, n$ has M_i repeated measurements collected at time k, $k = 1, ..., M_i$. For the purpose of this study, we set $M_i = M$. Assume that at each time point $k = 1, 2, \ldots, M$, we observe a time-dependent exposure T_{ik} taking values of $T_{ik} = 0$ or 1. We assume that exposure is on-going and is not discontinued, hence, once $T_{ik} = 1$, then $T_{iM} = 1 \; \forall M > k$. Let us also assume that for subject i we observe time-dependent confounders X_{ik} that may possibly be affected by previous exposures. For each subject i we denote the exposure and covariate history up to time k by $\bar{T}_{ik} = (T_{i1}, T_{i2}, \ldots, T_{iM})$ and $\bar{X}_{ik} = (X_{i1}, X_{i2}, \ldots, X_{iM})$. We let Y_{ik} be the observed outcome at time k, and subject i has a vector of repeated outcomes as $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{iM})^T$. For repeated measures, sometimes past outcomes may act as confounders on future outcomes and exposures are included in $\bar{X}_k$.

Potential outcomes are defined for each subject at the k time points to characterize the causal effects. Robins (1986) generalized the potential outcome framework of Rubin (1974) for longitudinal data. The potential outcome for subject i at time k for the entire treatment history $\bar{t}_k$ is given as $Y_{ik}(\bar{t}_k)$ denoting subject i's outcome Y at time k, had the subject been exposed $\bar{t}_k$, contrary to the truth. Where $\bar{t}_k$ is the potential treatment history. The observed outcome $Y_{ik} \equiv Y(\bar{T}_{ik})$ when $\bar{t}_k$ is equal to the actual exposure received.

The estimate of interest is the difference in the means of the potential outcomes $E[Y_k(\bar{t}_k)]$ between different exposure groups, $\bar{t}_k$, at time k

$$E[Y_{i,k}(\bar{t})] - E[Y_{i,k}(\bar{0})] \tag{3.1}$$

The average causal effect for repeated measurements can be written in vector notation as;

$$E[\vec{\mathbf{Y}}(\bar{t})] - E[\vec{\mathbf{Y}}(\bar{0})] \tag{3.2}$$

3.3 Marginal Structural Models (MSMs)

This section provides a detailed description of the MSM and the inverse probability treatment weights used to estimate the causal effects for repeated measures. The marginal structural model is also known as the model for potential (counterfactual) outcomes (Robins, 2000) model the mean of the counterfactual outcomes for a given treatment history. Marginal structural models (MSMs) are causal models designed to adjust for time-dependent confounding in observational studies with time-varying treatments. In assessing the effectiveness of a time-varying exposure, the MSM is useful in dealing with confounding by confounders affected by past exposure (treatment-confounder feedback) (Robins, 1997). The parameters of an MSM can be consistently estimated by the use of inverse probability weights (IPW) that create a pseudo-population in which the distribution of confounders is balanced across the treated and untreated groups.

The marginal structural model is links the mean potential outcome at time k to the function

of the exposure throughout time k. The model is defined as follows;

$$E[Y_{i,k+1}(\bar{t})] = g(h(\bar{t}_{i,k}); \boldsymbol{\beta}) \tag{3.3}$$

In matrix notation, equation 3.3 can be written as;

$$\begin{pmatrix} E[Y_{i1}(\bar{t}_1)] = g(h(\bar{t}_{i,0}); \boldsymbol{\beta}) \\ E[Y_{i2}(\bar{t}_2)] = g(h(\bar{t}_{i,1}); \boldsymbol{\beta}) \\ \vdots \\ E[Y_{iM}(\bar{t}_M)] = g(h(\bar{t}_{i,M}); \boldsymbol{\beta}) \end{pmatrix} \tag{3.4}$$

Where $h(\bar{t})$ is the function of the exposure history. The function relating the exposure history to the mean potential outcome is given as $g(\bar{t}_k; \boldsymbol{\beta})$. Where common functions of $g(\bar{t}_k; \boldsymbol{\beta})$ can be additive, multiplicative, exposure differing at different time points or a specific exposure at a time point k (Hernan et al., 2002). For example, $g(\bar{t}_k; \boldsymbol{\beta}) = \beta_0 + \beta_1 t_k + \beta_2 k$, $g(\bar{t}_k; \boldsymbol{\beta}) = \beta_0 + \beta_1 t_k + \beta_2 t_{k-1} + \beta_3 k$, $g(\bar{t}_k; \boldsymbol{\beta}) = \sum_{t=0}^{M} \beta_k t_k$, $g(\bar{t}_k; \boldsymbol{\beta}) = \beta_0 + \beta_1 \text{cum}(\bar{t}_k) + \beta_2 k$, where $\text{cum}(\bar{t}_k) = \sum_{k=1}^{K} t_k$. If the repeated measures Y_{ik} are continuous then equation 3.3 can be written as

$$E[Y_{i(k+1)}(\bar{t})] = \beta_1 \text{cum}(\bar{t}_k)$$

For binary $Y_{i(k+1)}$, equation 3.3 is given as

$$\text{logit}(\Pr(Y_{i(k+1)}(\bar{t}) = 1)) = \beta_1 \text{cum}(\bar{t}_k)$$

The average causal effect for equation 3.3 can be defined as $E[\mathbf{Y}_k(\bar{t})] - E[\mathbf{Y}_k(\bar{0})]$. Therefore, β_1 measures the change in the mean outcome caused by addition of the exposure.

The observed conditional expectation of the repeated measured outcomes $E[Y_{ik}|\bar{T}_k = \bar{t}_k, \mathbf{X}] \neq E[Y_{ik(\bar{t}_k)}]$, due to confounding by $\bar{X}_{ik}$. Even if we adjusted for the observed covariates for the observed conditional expectation, confounder bias would still be present. Robins (1999) showed that using the inverse probability of treatment weights, we could estimate the causal effect β_1. A weighted conditional expectation can be used to estimate the causal effect of the MSM. Robins (1999) showed that with the weights $E[\mathbf{1}(\bar{T}_i = \bar{t}_i)W_iY_i] = g(\bar{t}; \boldsymbol{\beta})$, where W_i is the weight and will be defined in the subsequent sections. In the following sections, we describe the assumptions needed to assess the causal effect of MSM using observational data and inverse probability treatment weights used to estimate the MSM parameters.

3.3.1 Assumptions for causal identification

In this section, we describe assumptions required to identify the causal effect $\boldsymbol{\beta}$ from the observed data. The assumptions are a modification of the strongly ignorable assumptions presented in chapter 1, section (1.4.1) to account for the time-varying nature of the outcomes, exposures, and confounders and are known as sequential ignorability assumptions. The following are sufficient assumptions to identify the causal effect from the MSM for repeated measures observational data.

Assumption 1: Consistency; $T_k = t_k$ implies $Y_{ik} \equiv Y_{ik}(\bar{T}_k)$

Assumption 2: Positivity; $0 < \mathbb{P}(T_k = t_k|\bar{T}_{k-1} = \bar{t}_{k-1}, \mathbf{X}_k = \mathbf{x}_k) < 1$

Assumption 3: Sequential exchangeability; $Y_k(\bar{t}) \amalg \bar{T}_{k-1}|\bar{\mathbf{X}}_{k-1}, \mathbf{X}_0$. The assumption states

that at each time point the exposure is conditionally independent of the potential outcomes.

Exchangeability assumes that the potential outcomes (the hypothetical outcomes under given exposure levels) are independent of the actual exposure level. For example, in assessing the effect of awareness of HIV positivity on condom use, then exchangeability would imply that the potential outcomes are condom use were an individual to be aware of HIV negativity, and condom use were that same individual to be aware of HIV positivity (two of these potential outcomes will be counterfactual). Exchangeability would mean that these potential outcomes are independent of an individual's actual exposure after accounting for all the measured confounders, including the exposure history. The positivity assumption implies that individuals have a non-zero probability of being either in exposure groups, for all the identified covariates and past exposure history.

3.3.2 Inverse probability treatment weights

The inverse probability of treatment weights (IPTW) is used to control for confounding. The IPTW is the inverse of the conditional probability of exposure at a specific time t, given prior exposure status and an individuals covariate history. The IPTW are calculated for each individual by first estimating the probability, $p = \mathrm{P}(T_k = 1|\bar{T}_{k-1}, \bar{\mathbf{X}})$. Under the assumptions presented in section 3.3.1, Robins (1999) showed that the inverse probability of treatment weighting (IPTW) could consistently estimate β_1 of the MSM in equation 3.3. The IPTW is the inverse of the conditional probability of the treatment at time k given the past treatment history, confounder history, and the baseline confounders. For subject i, the weight under treatment history $\bar{t}_k$, the IPTW are calculated by first estimating the probability, $P(\mathrm{T}_k = 1|\bar{T}_{k-1}, \bar{\mathbf{X}})$ using a multivariable logistic regression model and then

obtaining a product of their inverses to get the overall weight;

$$W_{ik} = \prod_{k=1}^{M} \frac{1}{P(T_k = 1|\bar{T}_{k-1} = \bar{t}_{k-1}, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k, \mathbf{X}_1)} \tag{3.5}$$

Individuals with a small probability of being exposed may obtain a large weight and vice versa resulting in extreme weights. Extreme weights affect the estimated effects by inflating the variance and confidence intervals (Chesnaye et al., 2021; Cole et al., 2003). Thus, extreme weights can be handled through either weight stabilisation or weight truncation (trimming). Stabilising the weights is achieved by replacing the numerator in (3.5) (which is 1 in the unstabilized weights) with the probability of exposure given past exposure and all the baseline confounders (i.e. without time-dependent covariates). The sIPTW are calculated as follows;

$$SW_{ik}^{t} = \prod_{k=1}^{M} \frac{P(T_k = 1|\bar{T}_{k-1} = \bar{t}_{k-1}, \mathbf{X}_0)}{P(T_k = 1|\bar{T}_{k-1} = \bar{t}_{k-1}, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k, \mathbf{X}_1)} \tag{3.6}$$

For observational data, the probability of being treated at time k given past treatment and the confounder history is unknown and must be estimated. Parametric and non-parametric approaches have been proposed for estimation. However, the most common parametric model used is the multivariate logistic regression model applied to each time point independently (Hernan, Brumback, & Robins, 2000)

$$P(T_{ik} = 1|\bar{T}_{i,k-1} = \bar{t}_{i,k-1}, \bar{\mathbf{X}}_{ik}) = \prod_{k=1}^{M} \{1 + \exp(-[(T_{i,k-1})\boldsymbol{\beta}_k, \bar{\mathbf{X}}_{ik}])\}$$

where $\boldsymbol{\beta}_k$ is the unknown coefficient to be estimated which include the causal parameters. The weights described in (3.5) and (3.6) only control for time-varying confounders and exposures and are not designed to account for loss-to-follow-up (LTFU). Thus, to control

for loss-to-follow-up, censoring weights are defined. Let D_t indicate loss-to-follow-up throughout the study, then $D_t = (1 - \text{censored and}, 0 - \text{not censored})$. We denote $\bar{D}_t = (D_0, D_1, \ldots, D_M)$ as the LTFU history. We assume no censoring at baseline ($D(0) = 0$). The weights for censoring were obtained by estimating the probability of remaining in the study (uncensored) at time $t + 1$ conditional on being uncensored in the past (t) and the observed confounders. The weights for LTFU are given as;

$$SW_{ik}^{d} = \prod_{k=1}^{M} \frac{P(D_{k+1} = 0 | \bar{D}_k = 0, \bar{T}_k = \bar{t}_k)}{P(D_{k+1} = 0 | \bar{D}_k = 0, \bar{T}_k = \bar{t}_k, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k)} \tag{3.7}$$

The final weight for subject i is calculated as a product of the stabilized inverse probability weights and the censoring weights at each time point k.

$$SW_{ik} = SW_{ik}^{t} \times SW_{ik}^{(}d) \tag{3.8}$$

3.3.3 Estimation

Under the assumptions stated in section 3.3.1, Robins, Hernan, and Brumback (2000) showed that a consistent estimator of $\boldsymbol{\beta}$ is a solution to the estimating equations of the form;

$$\sum_{i=1}^{n} D_i^T(\boldsymbol{\beta}) \mathbf{V}^{-1} SW_{ik}(Y_i - \mu_i(\boldsymbol{\beta})) = 0 \tag{3.9}$$

which is basically a weighted generalized estimating equation (GEE) with time-specific weights SW_{ik}. For $Y_i = (Y_{i1}, \ldots, Y_{iK})$, $D_i(\boldsymbol{\beta}) = \frac{\partial \mu_i(\beta)}{\partial \beta}$, $\mu_i(\boldsymbol{\beta}) = (g(T_{i1}, \mathbf{X}, \beta), \ldots,$ $g(\bar{T}_{iM}, \mathbf{X}, \beta))^T$, SW_i is the diagonal matrix of weights SW_{ik}. and $\mathbf{V}^{-1}$is the working

covariance matrix as

$$\mathbf{V} = \mathbf{A}^{\frac{1}{2}}\mathbf{R}(\boldsymbol{\alpha})\mathbf{A}^{\frac{1}{2}} \tag{3.10}$$

where $\mathbf{R}(\boldsymbol{\alpha})$ is a $M \times M$ working correlation matrix fully specified by the unknown correlation parameter $\boldsymbol{\alpha}$. For repeated measurements, the standard correlation structures for $\mathbf{R}(\boldsymbol{\alpha})$ can be assumed as exchangeable, unstructured and autoregressive with order one (AR(1)). For the exchangeable correlation structure we have

$$\text{Corr}(y_{ik}, y_{ik'}) = \alpha \forall,\ k \neq k^{'} \text{ then}$$
$$\mathbf{R} = \begin{cases} 1, & k = k^{'} \\ \alpha, & k \neq k^{'} \end{cases}$$

For the unstructured correlation structure we have

$$\text{Corr}(y_{ik}, y_{ik'}) = \alpha_{kk'},\ \forall k \neq k^{'} \text{ then}$$
$$\mathbf{R} = \begin{cases} 1, & k = k^{'} \\ \alpha_{kk'}, & k \neq k^{'} \end{cases}$$

For the auto-regressive correlation structure of order one we have

$$\text{Corr}(y_{ik}, y_{ik'}) = \alpha^{|k-k^{'}|},\ \forall k \neq k^{'} \text{ then}$$
$$\mathbf{R} = \begin{cases} 1, & k = k^{'} \\ \alpha^{|k-k^{'}|}, & k \neq k^{'} \end{cases}$$

A convenient choice for the working correlation structure is the working independence correlation (Hernan et al., 2002). However, one may also consider applying the exchange-

able working correlation, and AR(1) working correlation between repeated measurements within a subject.

It can be shown that the inverse probability weighting estimates the causal parameters of the marginal structural model. For example, for exposure measured at two time points T_1 and T_2, $E[Y_{t_1,t_2}]$ is equal to the expectation of the observed outcome Y conditional on past treatment and covariate history. Following (Naimi, Cole, & Kennedy, 2017) we have;

$$E[Y_{t_1,t_2}] = E\left[\frac{I(T_1 = t_1, T_2 = t_2)Y}{P(T_2 = t_2|\mathbf{X}, T_1 = t_1)P(T_1 = t_1)}\right]$$

$$E\left[E\left(\frac{I(T_1 = t_1, T_2 = t_2)Y}{P(T_2 = t_2|\mathbf{X}, T_1 = t_1)P(T_1 = t_1)}|T_1 = t_1, \mathbf{X}, T_2 = t_2\right)\right]$$

$$E\left[\frac{I(T_1 = t_1, T_2 = t_2)E(Y|T_1 = t_1, \mathbf{X}, T_1 = t_1)}{P(T_2 = t_2|\mathbf{X}, T_1 = t_1)P(T_1 = t_1)}\right]$$

$$E\left[\frac{I(T_1 = t_1)E(Y|T_1 = t_1, \mathbf{X}, T_2 = t_2)}{P(T_1 = t_1)}\right]$$

$$E\left[\frac{I(T_1 = t_1)E\left[E(Y|T_1 = t_1, \mathbf{X}, T_2 = t_2)|T_1 = t_1\right]}{P(T_1 = t_1)}\right]$$

$$E\left[E(Y|T_1 = t_1, \mathbf{X}, T_2 = t_2)|T_1 = t_1\right]$$

$$E[Y_{t_1,t_2}] = E\left[E(Y|T_1 = t_1, \mathbf{X}, T_2 = t_2)\right] \tag{3.11}$$

The proofs falls under the consistency $Y_k(\bar{t}) = Y_k$ for $k = 1, 2$, iterated expectations, positivity and exchangeability assumptions.

When $g(.)$ of equation 3.3 is the identity link function then the equation 3.9 reduces to the weighted least squares. Robins (1999) showed that for the stabilized inverse probability weights 3.6, the equation is consistent for

$$\mathrm{E}[U(\boldsymbol{\beta})] = 0$$

$$\mathrm{E}\left[\sum_{i=1}^{n} D_i^T(\boldsymbol{\beta})\mathbf{V}^{-1}SW_i(Y_i - \mu_i(\boldsymbol{\beta}))\right] = 0$$

3.4 Motivating Example

Risky sexual behaviour such as premarital sex, multiple sexual partners, and unprotected sex has been shown to be major contributors to new HIV infections and account for the high HIV transmission in sub-Saharan Africa (Machira, Maonga, & Chirwa, 2020; UNAIDS, 2015; Wondemagegn & Berkessa, 2020; Ortblad et al., 2019). With the increase in programs that promote HIV testing and HIV status awareness, research interest is now shifting to assess the effect of HIV status awareness on risky sexual behaviour (Palk & Blower, 2018; Adaralegbe, Adeoti, Ogbonda, Egeolu, & Ukoha, 2021; Rosenberg et al., 2017). Most of the empirical evidence on the association between awareness of HIV status and risky sexual behaviour comes from using cross-sectional studies (Bunnell et al., 2008; Guiella & Madise, 2007; Huerga et al., 2017; Ventura-Filipe & Newman, 1998), which could be limited as the exposure and outcome are observed simultaneously. Longitudinal studies provide the temporal dependence between the exposure and outcome to estimate

the causal effect. Thus, we need awareness of HIV status (exposure) to precede sexual behaviour (outcome) to estimate the causal effect correctly. We used the MSM approach on data from the Malawi Longitudinal Study of Families and Health (MLSFH) to estimate the causal effect of HIV positivity awareness on condom use and multiple sexual partners.

3.4.1 Application on The Malawi Longitudinal Study of Families and Health (MLSFH)

The Malawi Longitudinal Study of Families and Health (MLSFH) which collects socio-economic, family dynamics and health information including HIV for up to 4000 couples in the 120 villages across the three districts were used (H.-P. Kohler et al., 2014; Baranov & Kohler, 2018; I. Kohler et al., 2020) . For this paper, 2540 individuals enrolled in 2004 and followed up at eight time points namely; 2006, 2008, 2010, 2012, 2013, 2017, 2018, and 2019 were studied.

3.4.2 Exposure variable

This study concentrated on individuals who were aware of their HIV test results. In each year, awareness of HIV positivity was defined as an individual's knowledge of their positive HIV test result. Over the entire period from 2004 to 2019, the cumulative measure of awareness of an individual's HIV positivity was measured as the number of times the individual was aware of their positive HIV test result.

3.4.3 Outcome variable(s)

In this study, risky sexual behaviour was measured by the number of sexual partners and condom use at last sex. Condom use at last sex was categorised as 1 (=Yes) of 0 (=No) and for number of sexual partners, it was 1 ($\geq$ 2) or 0 (if 0, 1)

3.4.3.1 Confounders

The following predictor variables were considered as baseline confounders: region, religion, wealth, given an incentive to collect HIV test results, distance to HIV testing and counselling (HTC) centres, perceived risk of HIV infection, gender, and comprehensive HIV knowledge. Age in years and education level (0 to 12 years) were considered as time-varying confounders. Figure 1 illustrates this relationship.

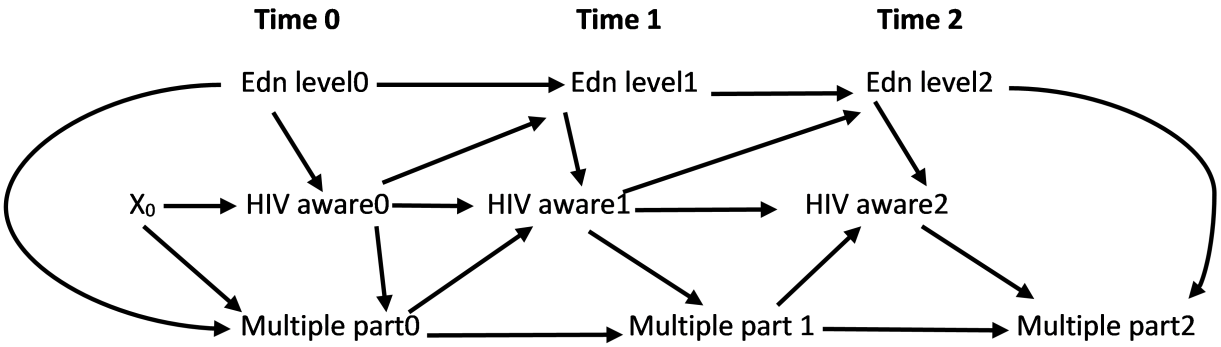


Figure 3.1: Directed Acyclic Graph (DAG) on HIV status awareness on multiple sexual partners for three time periods where X_0 denotes the baseline confounders. **Source**: Researcher (2023)

3.4.4 Description of the data

Table 3.1 presents a description of the data for the five MLSFH waves from 2004 to 2019. We present the HIV sample observed and the total sample reported by (H.-P. Kohler et al., 2014). The percentage of the sample is based on the total observed sample. In 2004, the total sample was 4094 for both ever married and never-married individuals. The total sample representing those who were tested for HIV was 2894 in the year 2004 resulting in a percentage of 70.7%. The percentage validity of those tested for HIV decreased across the years. The HIV data for the year 2010 comprise those who were followed from the year 2004 (original sample of 2004) and not exactly based on those who had an HIV test result, as HIV testing was not conducted in 2010 due to the low HIV incidence in the MLSFH data (H.-P. Kohler et al., 2014).

Table 3.1: Sample by wave for MLSFH participants

Wave	**HIV sample**	**Sample collected**
2004	2540	4094
2006	1682	4980
2008	1530	6376
2010	1500	6331
2012	664	1266
2013	663	1257
2017	893	1606
2018	893	1626
2019	729	2865

3.4.5 Estimation of HIV-status awareness effect

We modeled the effect of HIV positivity awareness on the outcomes over time k using the following model;

$$\text{logit}P(Y_{k+1} = 1|\bar{T}_k = \bar{t}_k) = \beta_0 + \beta_1 \sum_{k=1}^{9} A_t \tag{3.12}$$

Where β_0 is the unknown intercept, β_1 is the estimated effect of cumulative awareness of HIV positivity ($\sum_{k=1}^{9} A_t$) compared to HIV negativity awareness on reducing multiple sexual partners or increase condom use over time, as a hypothesis. Equation 3.12 assumes that each additional period of awareness of HIV status has an additive effect (the effect over time) on change in condom use and multiple sexual partners

The effects of HIV positivity awareness on condom use and having multiple sexual partners were estimated using a weighted (IPTW) marginal model (Hernan et al., 2002). With the weights, the distribution of the confounders of the analytical sample across the known HIV-status awareness groups were comparable, mimicking a randomized control trial. The weights were calculated using the `IPW` package in R (van der Wal, Geskus, et al., 2011) and the outcome models were fitted using the `geepack` package in R (Højsgaard, Halekoh, & Yan, 2006). We considered an AR(1) working correlation structure, $R_i(\alpha) = \alpha^{|k-k'|}$ to estimate the causal parameters. Robust standard errors were calculated using the sandwich variance estimator to account for the repeated measurements. Studies have shown that the exchangeable correlation structures produce biased estimates (Keogh, Daniel, Vanderweele, & Vansteelandt, 2017; Sullivan Pepe & Anderson, 1994). This mainly occurs when the past outcome affects future exposures which is the case for our example

as shown in Figure (3.1). The weighted models were compared to the unweighted model to assess the potential impact of time-dependent confounding on the estimated effect of HIV-status awareness on the outcomes.

The stabilised inverse probability treatment weights (sIPTW) were used to control for confounding. The sIPTW were calculated as

$$SW_i^a(t) = \prod_{k=1}^{9} \frac{P(A_k = 1|\bar{A}_{k-1} = \bar{a}_{k-1}, \mathbf{X}_1)}{P(A_k = 1|\bar{A}_{k-1} = \bar{a}_{k-1}, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k, \mathbf{X}_1)} \tag{3.13}$$

We calculated censoring weights by considering individuals who refused to participate, were dead, moved out or not eligible to participate in the MLSFH study as censored while individuals who were temporarily absent were uncensored. Denote $\bar{D}_t = (D_1, D_2, \ldots, D_9)$ as the loss-to-follow-up history with the assumption of no censoring at baseline ($D(1) = 0$), the censoring weights were calculated as follows

$$SW_i^d(t) = \prod_{k=1}^{0} \frac{P(D_k = 0|\bar{D}_{k-1} = 0, \bar{A}_k = \bar{a}_k)}{P(D_k = 0|\bar{D}_{k-1} = 0, \bar{A}_k = \bar{a}_k, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k)} \tag{3.14}$$

The final weights used in estimating the causal parameters was calculated as a product of the stabilised and censoring weights.

3.4.6 Sensitivity Analysis

The estimation of the causal effect of model 3.12 is done under the strongly ignorable assumption (particularly, no unmeasured confounder assumption) in estimating the causal effect of HIV positivity awareness on condom use and multiple sexual partners. However,

the assumption of no unmeasured confounders cannot be tested, and the IPTW is limited in that it cannot control for the unmeasured confounders. There is a possibility that the unobserved confounders are associated with the exposure, and the outcomes which may bias the estimated effects. The magnitude of bias due to unmeasured confounding may vary and can be either large or small. If the bias is small, the estimates may not change due to unmeasured confounders. However, if the bias is significant, the estimates may be affected. Hypothetical confounders that may have been measured such as disclosure of a partner's status, knowledge of a partner's HIV status, attitudes towards condom use may be associated with condom use, and alcohol use may be associated with multiple sexual partners. Therefore, an important approach of evaluating for causation when dealing with unmeasured confounding is to conduct a sensitivity analysis. There are several approaches of conducting a sensitivity analysis, either using a Mantel Haenszel test or McNemars test (for matched pairs) to identify the strength of association between the unobserved confounders and the exposure that would render the causal association non-causal (Rosenbaum, 2002). Another sensitivity analysis measure is known as the E-value which implies "evidence for causality" (VanderWeele & Ding, 2017; Mathur, Ding, Riddell, & VanderWeele, 2018). The E-value measures the minimum association of unmeasured confounders on the observed exposure-outcome relationship. It measures how strong a confounding factor would have to be to explain away the observed association between exposure and outcome. The higher the E-value, the greater the magnitude of the unmeasured confounder required to render the estimated treatment exposure effects non-significant. The E-value increases as the strength of the unmeasured confounder(s) that could explain away the observed association increases, indicating that a higher E-value is associated with a more robust association.

We assessed the strength of association between the unmeasured confounder(s) and the exposure and the outcome that would shift the observed causal effect towards the null (non-causal association) using the E-value approach.

3.5 Results

The distribution of awareness of HIV status by year of collection is shown in Table 3.2. A majority (92.8%) of the individuals were HIV negative while 183 (7.2%) were HIV positive. In 2004, more than half of those who were HIV positive (62.8%) were aware of their HIV positivity while 68 (37.2%) were unaware of their HIV positive status. Awareness of HIV positivity increased over time among those who were HIV positive, with a majority of individuals in 2006 (96.8%) and 2008 (97.4%) being aware of their HIV positivity. All the individuals who were HIV positive in 2010 and 2012 were aware of their HIV positive status. In 2017 and 2018, 73 individuals who were HIV positive were aware of their HIV positivity status. In subsequent years from 2004, awareness of HIV positivity increased.

Table 3.2: Distribution of HIV status awareness by wave

	Awareness of HIV status						
	HIV positivity (N=183 (7.2%))			HIV negativity (N=2357 (92.8%))			
	Yes	No	Subtotal	Yes	No	Subtotal	Total
Year							
2004	115 (62.8%)	68 (37.2%)	183	1626 (69.0%)	731 (31.0%)	2357	2540
2006	91 (96.8%)	3 (3.2%)	94	1568 (98.7%)	20 (1.3%)	1588	1682
2008	74 (97.4%)	2 (2.6%)	76	1448 (99.6%)	6 (0.4%)	1454	1530
2010	74 (97.4%)	2 (2.6%)	76	1418 (99.6%)	6 (0.4%)	1424	1500
2012	42 (100%)	0 (0.0%)	42	607 (97.6%)	15 (2.4%)	612	664
2013	42 (100%)	0 (0.0%)	42	606 (97.6%)	15 (2.4%)	621	663
2017	73 (98.6%)	1 (1.4%)	74	819 (100%)	0 (0.0%)	819	893
2018	73 (98.6%)	1 (1.4%)	74	819 (100%)	0 (0.0%)	819	893
2019	58 (100%)	0 (0.0%)	58	671 (100%)	0 (0.0%)	671	729

Table 3.3 presents the distribution of baseline characteristics for the individuals by awareness of HIV status and their respective associations. Out of the 115 individuals aware of their HIV positivity status, 67 (58.3%) were females, 48 (41.7%) were males, and 83 (72.2%) were aged 15-44 years. A majority of the individuals where from the central region of Malawi (44.4%), and 76.5% had some education. A majority of those who were not aware of their positive HIV status were female (61.8%), were from medium or rich households (69.6%), resided in the central region (31 (45.6%)) of Malawi, and were Christian (48.6%). Religion, having received an incentive, and region were all significantly associated with awareness of HIV status. We also observed an association between HIV status awareness and condom use and multiple sexual partners. At baseline, among

individuals who were aware of their positive HIV status, 26.1% had used a condom during their last sex, and 80.8% had multiple sexual partners.

The study had individual dropouts either due to death, migration, or not being eligible to participate. Thus, the distribution of the purported confounders differed by loss-to-follow-up (LTFU) for the succeeding year of 2004 was assessed, and the findings are presented in Table 3.4. The findings show differences in distribution between the LTFU levels among some of the measured confounders; hence, we controlled for loss-to-follow-up in the analysis.

Table 3.3: Distribution of baseline characteristics by HIV status awareness of study participants in the MLSFH for year 2004

	Awareness of HIV status		
	HIV positivity	**HIV negativity**	**p-value**
Total (n=1741)	115 (6.6%)	1626 (93.4%)	
Gender			
Women	67 (58.3)	870 (53.5%)	
Men	48 (41.7%)	756 (46.5%)	0.323
Age category			
15-44 years	83 (72.2%)	1182 (72.7%)	
45 years above	32 (27.8%)	444 (27.3%)	0.525
Wealth			
Poor	45 (45%)	547 (41.9%)	
Medium/rich	55 (55.0%)	759 (58.1%)	0.543
Education			
None	23 (23.5%)	355 (24.8%)	
Some education	75 (76.5%)	1076 (75.2%)	0.766
Religion			
Christian	53 (48.6%)	859 (55.4%)	
Muslim or other	56 (51.4%)	677 (44.1%)	0.043
Incentive			
No	34 (45.9%)	482 (48.5%)	
Yes	40 (54.1%)	512 (51.5%)	0.044
Region			
North	43 (37.4%)	536 (32.9%)	
Central	51 (44.4%)	600 (36.9%)	
South	21 (18.3%)	490 (30.2%)	0.025
Condom Use			
No	68 (73.9%)	1001 (74.5%)	
Yes	24 (26.1%)	342 (25.5%)	0.895
Multiple partners			
No	20 (19.2%)	535 (39.1%)	
Yes	84 (80.8%)	832 (60.9%)	$<$0.001

Table 3.4: Distribution of characteristics by loss-to-follow-up in the MLSFH in 2006

	Loss-to-follow-up (LTFU)		
	Yes	**No**	**p-value**
HIV awareness			
HIV positivity	72 (90.0)	1306 (94.8%)	
HIV-negativity	8 (10.0%)	72 (5.2%)	0.068
Gender			
Female	54 (67.5)	905 (65.7%)	
Male	26 (32.5%)	474 (34.3%)	0.738
Age category			
14-44 years	61 (76.3%)	964 (69.9%)	
45 years above	19 (23.8%)	414 (30.0%)	0.231
Wealth			
Poor	24 (30%)	503 (36.5%)	
Medium/rich	56 (70.0%)	875 (63.5%)	0.239
Education			
None	15 (18.8%)	321 (23.3%)	
Primary	57 (71.3%)	912 (66.2%)	
Secondary/Tertiary	8 (10.0%)	145 (10.5%)	0.610
Religion			
Christian	47 (58.8%)	789 (57.3%)	
Muslim or other	33 (41.3%)	589 (42.7%)	0.097
Region			
North	44 (55.0%)	281 (20.4%)	
Central	17 (21.3%)	531 (38.5%)	
South	19 (23.8%)	566 (41.1%)	$<$0.001

The distributions of the weights used for the MSM are shown in Figure 3.2 (Distribution of the log of non-stabilized and stabilized weights) for several follow-up times. The weights are presented as a logarithmic form for display purposes. The unstabilized weights produced a skewed distribution with extreme weights.

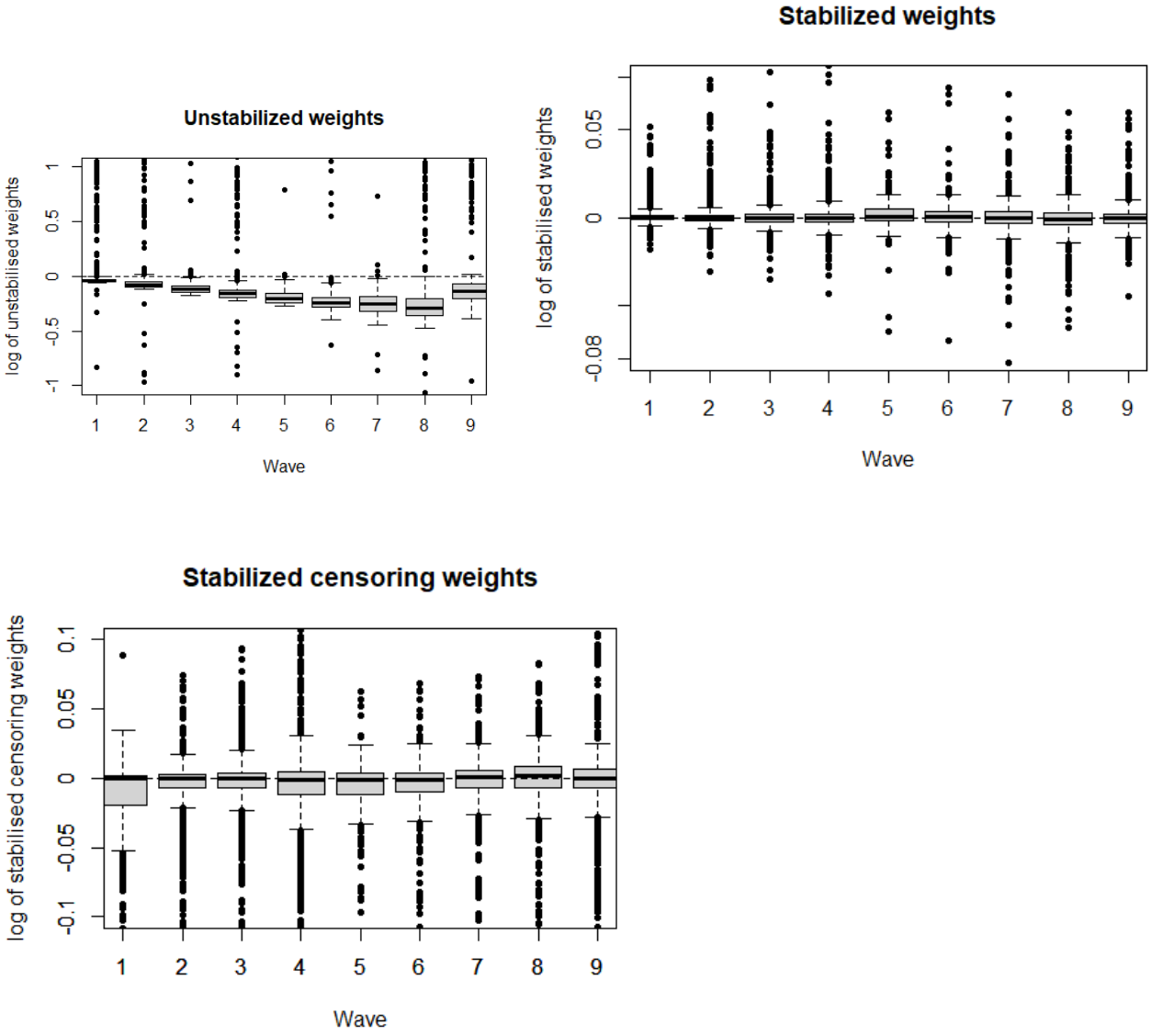


Figure 3.2: Distribution of the log of unstabilized, stabilized and censoring weights. Source: Researcher(2023)

We obtained the IPTWs by taking the product of the inverse probability treatment weights at each time point. The final weights was a product of the stabilized weights and the censoring weights. Table 3.5 presents the distribution of the unstabilised, stabilized and loss-to-follow-up weights. The unstabilised weights had a mean of 0.89, a median of 0.88, a standard deviation of 0.5, and an inter-quartile range of (0.81, 0.95). The maximum weight was 16.4, implying that an individual with the largest weight contributed about 16 people to the pseudo-population. The unstabilized weights had a huge variability across the time points, see Figure 3.2. Conversely, the stabilized weights had a median and mean centred at 1.0 across all time points, as shown in Figure 2b. The range narrowed to a minimum of 0.99 and a maximum of 1.00. The maximum weight was 1.2, which corresponded to a contribution of 1 person to the pseudo-population. The LTFU weights had a mean of 0.99 and median of 0.99 and ranged between 0.95 and 1.03, see Figure 3.2. The stabilized weights were less variable (SD=0.01) and its mean and median were centered around 1.0 as required, compared to the LTFU and the unstabilised weights.

Table 3.5: Distribution of inverse probability weights

Weights	**Mean (SD)**	**Median**	**Interquartile range**	**Min**	**Max**
Unstabilized	0.89 (0.5)	0.88	(0.81, 0.95)	0.0	16.40
Stabilized	1.0 (0.01)	1.0	(0.99, 1.00)	0.8	1.2
Censoring	0.99 (0.18)	0.99	(0.95, 1.03)	0.16	5.97

3.5.1 Confounder balance after weighting

From Figures 3.3 and 3.4, some the baseline confounders such as region, age and incentive were not balanced between HIV status awareness. After weighting, the standardized

difference for the confounders were closer to zero, signifying balance of the confounders between the exposure groups.

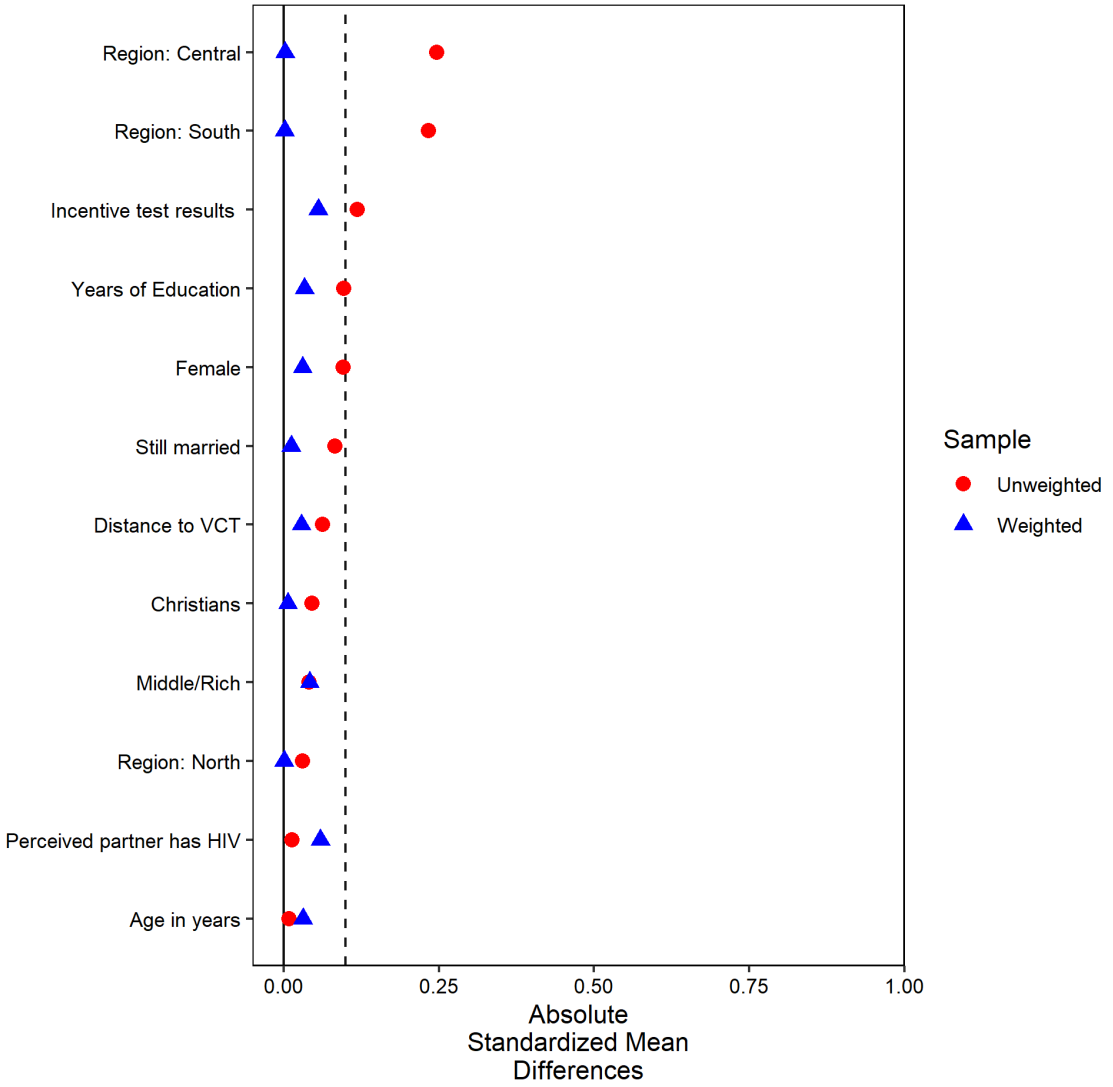


Figure 3.3: Covariate balance at baseline. **Source**: Researcher (2023)

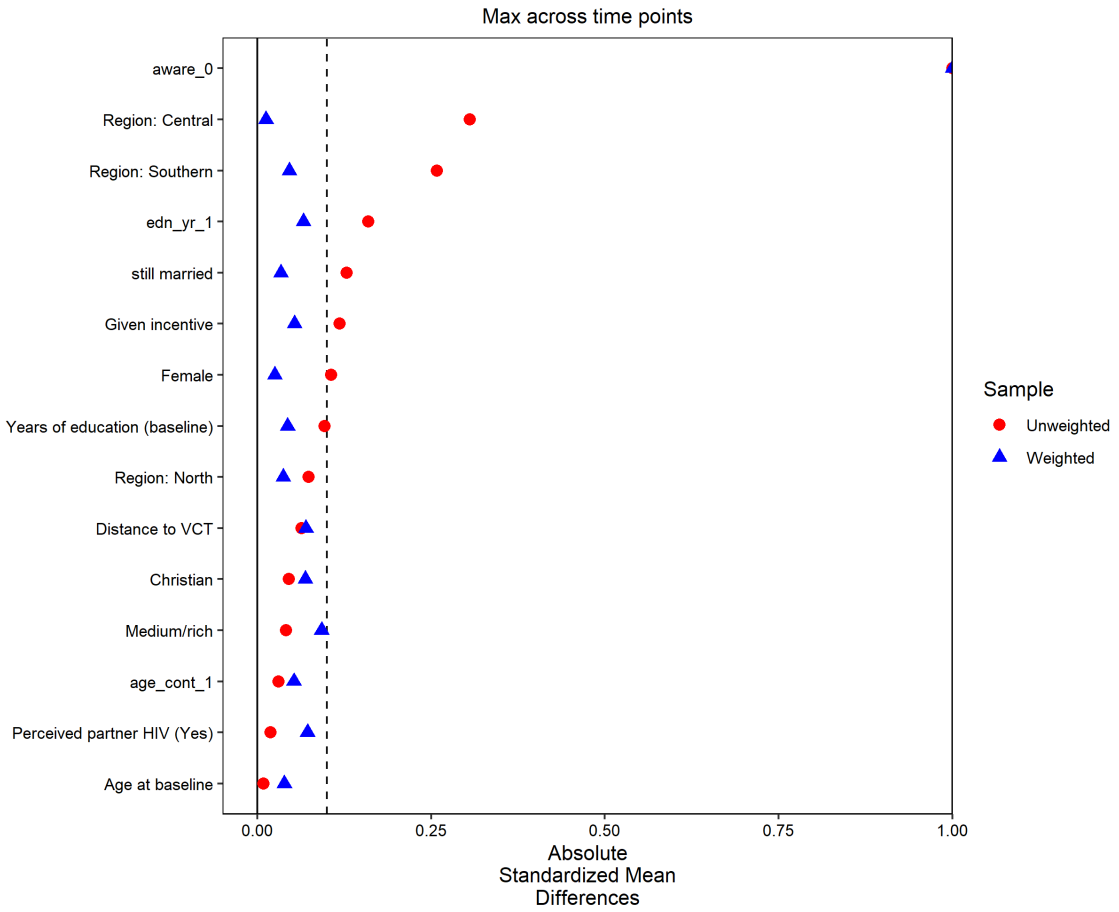


Figure 3.4: Confounder balance across two time points [1]1. **Source**: Researcher (2023)

3.5.2 Effect of HIV status awareness on risky sexual behaviour

A weighted (IPTW) marginal logistic regression model was used to estimate the awareness of HIV positivity effects. The weights were calculated using the `IPW` package (van der Wal et al., 2011). For modelling repeated multiple partners and condom use measurements, independent and AR(1) working correlation matrices were assumed and we estimated the variance using "robust" standard errors using the the `geepack` package in R (Højsgaard et al., 2006). An independent working correlation matrix was used in our analyses because

[1]aware_0 - Awareness at baseline, edn_yr_1 - Years of Education (2006),age_cont_1 - Age in years (2006)

other correlation structures defined using the generalized estimating equations (GEEs) have been shown to produce biased treatment effect estimates when past outcomes affect future exposures (in our example having multiple sexual partners may influence an individuals decision to know their HIV positivity) (Keogh et al., 2017; Sullivan Pepe & Anderson, 1994). The weighted models were compared to the estimated treatment effects obtained from the standard multivariable binary logistic model were we also adjusted for the effect of year.

3.5.2.1 Standard multivariable logistic regression model

Table 3.6 presents the odds ratio for the association between awareness of HIV positivity and condom use or multiple sexual partners. Condom use was significantly associated with awareness of HIV positivity (OR: 2.30; 95% CI=(1.82, 2.90)), (Table 3.6a). Additionally, men (OR: 1.72; 95% CI=(1.50, 1.97)); individuals with a primary (OR: 1.32; 95% CI=(1.08, 1.62)) and secondary education (OR: 1.73; 95% CI=(1.31, 2.29)); individuals from medium or rich households (OR: 1.16; 95% CI=(1.02, 1.32)) and those from the southern region (OR: 2.12; 95% CI=(1.34, 3.37)) were more likely to use condoms.

Table 3.6b presents the association between multiple sexual partners and awareness of HIV positivity adjusting for the measured confounders. Individuals aware of HIV positivity were more likely to have multiple partners (OR: 1.52; 95% CI=(1.17, 1.98)). Men were more likely to have multiple partners than women (OR: 3.11; 95% CI=(2.67, 3.62)). On the other hand, individuals from medium or rich households (OR: 0.81; 95% CI=(0.71, 0.93)) were less likely to have multiple sexual partners. In addition, over time individuals were less likely to have multiple partners (OR: 0.71; 95% CI=(0.68, 0.75)).

Table 3.6: Effect estimates (Odds ratios, (95% confidence Interval)) using standard multivariable logistic model

(a) Condom Use

Covariates	**Odds ratio**	**95% Conf. Int.**
Intercept	1.62	(1.23, 2.14)
HIV aware (HIV negativity vs HIV positivity)	2.30	(1.82, 2.90)
Gender (women vs men)	1.72	(1.50, 1.97)
Age (in years)	0.43	(0.37, 0.50)
Wealth (poor vs medium/rich)	1.16	(1.02, 1.32)
Education (none vs primary)	1.32	(1.08, 1.62)
Education (none vs secondary/Tertiary)	1.73	(1.31, 2.29)
Religion (Muslim or other vs Christian)	1.02	(0.88, 1.18)
Region (North vs Central)	0.99	(0.74, 1.32)
Region (North vs South)	2.12	(1.34, 3.37)
Year	0.99	(0.96, 1.02)

(b) Multiple Sexual Partners

Covariates	**Odds ratio**	**95% Conf. Int.**
Intercept	0.31	(0.23, 0.42)
HIV aware (HIV-negativity vs HIV positivity)	1.52	(1.17, 1.98)
Gender (women vs men)	3.11	(2.67, 3.62)
Age (in years)	1.00	(0.99, 1.01)
Wealth (poor vs medium/rich)	0.81	(0.71, 0.93)
Education (none vs primary)	1.07	(0.89, 1.30)
Education (none vs secondary/Tertiary)	0.88	(0.62, 1.25)
Religion (Muslim or other vs Christian)	1.16	(0.98, 1.37)
Region (North vs Central	1.22	(0.90, 1.68)
Region (North vs South)	0.95	(0.56, 1.63)
Year	0.71	(0.68, 0.75)

95% Conf. Int. - 95% Confidence Interval

3.5.2.2 Marginal structural model

Both the unstabilised and stabilised weights were used to fit the MSM model. Only the results for the stabilised weights are reported (see, Table 3.7) as the distribution of the weights were more stable (the results for the unstabilised weights are reported in Table 1 in the Appendix). In using the MSM, the higher the number of times a person was aware of their HIV positivity resulted in an increase in condom use (OR: 2.22; 95% CI=(1.79, 2.75)) compared to those individuals aware of HIV negativity (see Table 3.7a). However, the cumulative effect of awareness of HIV positivity on multiple sexual partners was not significant (OR: 1.06; 95% CI=(0.83, 1.35)) (see Table 3.7b).

Based on the results, the standard logistic regression model showed an increased effect on multiple sexual partners. On the other hand, the MSM produced a decreased effect of the multiple sexual partner estimates. From the findings, the unstabilised weighted model produced treatment effect estimates with wider confidence intervals implying large standard errors due to the high variability in the weights, unlike the estimated treatment effects from the stabilized weighted model.

Table 3.7: Effect Estimates of awareness of HIV positivity on risky sexual behavior using stabilized weights

(a) Condom Use

	Odds ratio	95% CI	p-value
Intercept	0.11	(0.10, 0.14)	<0.001
HIV negativity	1.00	-	-
HIV positivity	2.22	(1.79, 2.75)	<0.001

(b) Multiple Sexual Partners

	Odds ratio	95% CI	p-value
Intercept	0.24	(0.21, 0.27)	<0.001
HIV negativity	1.00	-	-
HIV positivity	0.96	(0.83, 1.35)	0.63

-We controlled for the following confounders: gender; wealth; age; education; time (wave); religion, comprehensive HIV knowledge; distance to VCT; received an incentive; region.

-95% CI = 95% Confidence Interval

3.5.2.3 The effect of awareness of HIV positivity status by gender

The effect of awareness of HIV positivity on condom use and multiple sexual partners was further stratified by gender (Table 3.8). Both men (OR= 2.32; 95% CI:(1.66, 3.24)) and women (OR= 2.24; 95% CI:(1.70, 2.95)) with a higher number of times aware of their HIV positivity were more likely to use condoms (Table 3.8a). However, only the women aware of their HIV positivity were more likely to have multiple sexual partners (OR= 1.76; 95% CI:(1.36, 2.28)) (Table 3.8b).

Further assessing the effect of condom use on multiple sexual partners and vice versa stratified by gender showed that men (OR= 1.45; 95% CI:(1.19, 1.76)) and women (OR= 1.29; 95% CI:(1.02, 1.63)) who had multiple sexual partners were more likely to use condoms (Table 3.8a). Similarly, men (OR= 1.41; 95% CI:(1.18, 1.67)) and women (OR= 1.28; 95% CI:(1.01, 1.62)) who used condoms were more likely to have multiple partners (Table 3.8b). This may imply the possible correlation between the risky sexual behaviour outcomes.

Table 3.8: Effect estimates of awareness of HIV positivity by gender

(a) Condom Use

	Men		Women	
	Odds ratio	**95% CI**	**Odds ratio**	**95% CI**
Intercept	0.19	(0.14, 0.26)	0.10	(0.07, 0.11)
Awareness				
HIV negativity	1.00	-	1.00	-
HIV positivity	2.32	(1.66, 3.24)	2.24	(1.70, 2.95)
Multiple partners				
No	1.00	-	1.00	-
Yes	1.45	(1.19, 1.76)	1.29	(1.02, 1.63)

(b) Multiple Sexual Partners

	Men		Women	
	Odds ratio	**95% CI**	**Odds ratio**	**95% CI**
Intercept	0.60	(0.46, 0.77)	0.14	(0.12, 0.17)
Awareness				
HIV negativity	1.00	-	1.00	-
HIV positivity	0.76	(0.49, 1.18)	1.76	(1.36, 2.28)
Condom use				
No	1.00	-	1.00	-
Yes	1.41	(1.18, 1.67)	1.28	(1.01, 1.62)

95% CI - 95% Confidence Interval

3.6 The effect of HIV status awareness on marital status

In addition to HIV status awareness affecting risky sexual behaviour, it also affects marriage dissolution (Reniers & Armbruster, 2012; Anglewicz & Reniers, 2014; Fedor, Kohler, & Behrman, 2015). While marriage dissolution among HIV-infected individuals may reduce HIV transmission, women are more likely to suffer in such circumstances, particularly among those from poor households (Porter et al., 2004). Therefore, we further assessed the effect of HIV positivity awareness on marital status. Table 3.10 presents the results of the effects. Regardless of HIV status, individuals were less likely to remain married over time. Individuals who were aware of their HIV positivity status were 32% less likely to be still married than those who were aware of their HIV negativity status.

Table 3.9: Effect of HIV status awareness on marriage

Covariates	**Odds ratio**	**95% Conf. Int.**
Intercept	0.31	(0.23, 0.42)
HIV aware (HIV-negativity vs HIV positivity)	0.57	(0.43, 0.77)
Gender (women vs men)	1.43	(1.18, 1.72)
Age (in years)	1.00	(0.99, 1.01)
Wealth (poor vs medium/rich)	1.48	(1.30, 1.68)
Education (none vs primary)	1.07	(0.89, 1.30)
Education (none vs secondary/Tertiary)	0.88	(0.62, 1.25)
Religion (Muslim or other vs Christian)	1.16	(0.98, 1.37)
Region (North vs Central	1.17	(0.93, 1.47)
Region (North vs South)	1.26	(1.00, 1.59)
Year	0.71	(0.68, 0.75)

Table 3.10: The effect of knowledge of HIV status on marriage outcomes over time

		Married
Weight type	**HIV awareness**	**Odds ratio (95% CI)**
Unstabilized weights	HIV negativity	1.00
	HIV positivity	0.66 (0.36, 1.21)
Stabilized weights	HIV negativity	1.00
	HIV positivity	0.62 (0.46, 0.83)

Table 3.11 presents HIV status awareness effect on marital stability dis-aggregated by gender. Both men and women who were aware of their positive HIV status were less likely to remain married (OR:0.68; 95% CI:0.13, 3.58) and (OR: 0.23; 95% CI:0.13, 0.42), respectively, as compared to men and women who were not aware of their negative HIV status.

Table 3.11: HIV status awareness effect on marriage by gender

		Married
		Odds ratio (95% CI)
Males	**HIV negativity**	1.00
	HIV positivity	0.89 (0.45, 1.76)
Females	**HIV negativity**	1.00
	HIV positivity	0.53 (0.39, 0.73)

3.7 Sensitivity analysis

Table 3.12 present the calculated E-values for the estimates presented in Table 3.7. It was found that the effect on multiple sexual partners could have been affected greatly with a smaller impact of unmeasured confounders, but a larger impact could have been required to alter the effect on condom use. Therefore, the effect of awareness of HIV positivity on condom use was robust to unmeasured confounders While the estimated treatment effect on multiple sexual partners were more likely to be affected by unmeasured confounders.

Additionally, we conducted a sensitivity analysis to assess the influence of each of the confounders on the estimated HIV positivity awareness effect. The sensitivity analysis was done by calculating the weights while dropping some variables and obtaining the effect. From Figure 3.5, if all confounders were not included in the weights, then the effect would change in the direction of a no effect. However, not much change in the effect was observed if the individual variables were dropped.

Table 3.12: E-value parameters calculated as a function of the adjusted odds ratio (OR) for the effect of a positive HIV status awareness on the outcomes

	Condom use	Multiple sexual partners
Awareness	**Estimate (LL of CI)**	**Estimate (LL of CI)**
HIV positivity	3.87 (2.98;)	1.31 (1.00;)

LL of CI - Lower limit of confidence Interval

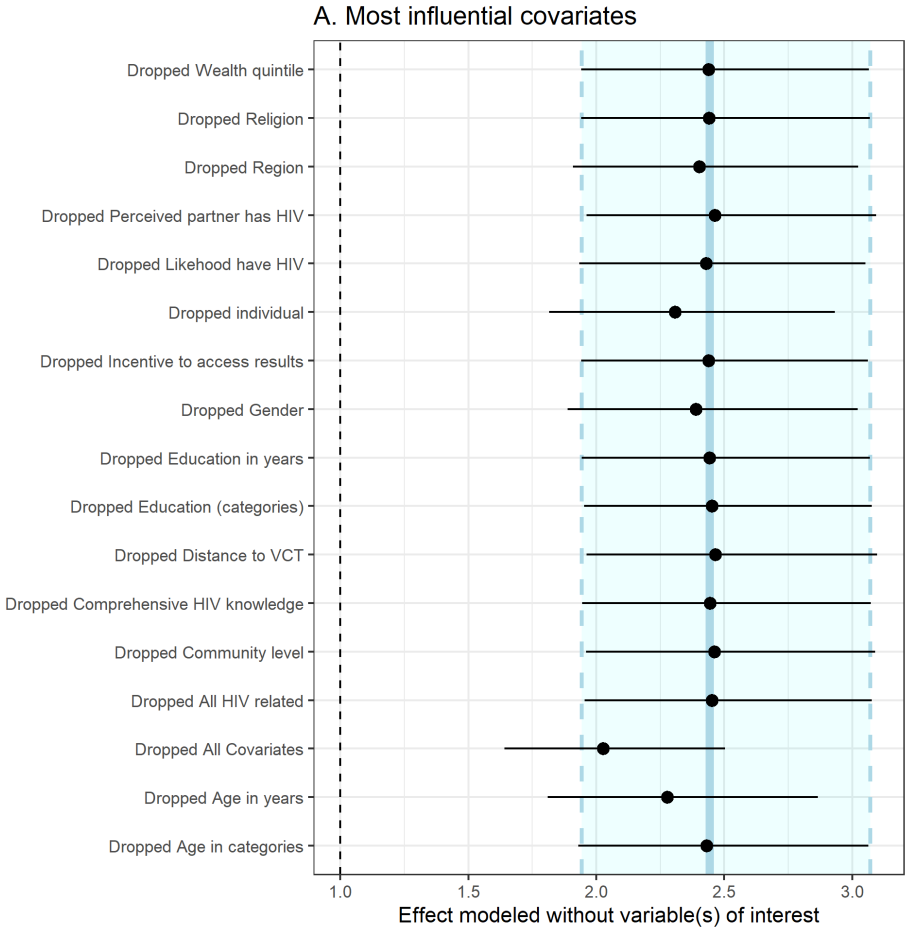


Figure 3.5: Influence of the confounders on the estimated effect. **Source:** Researcher (2023)

Overall from the results, we observe that the standard logistic model either underestimated or overestimated the effect of HIV status awareness on condom use and multiple sexual partners and marriage compared to estimates obtained from the model using the stabilized inverse probability weights. In addition, the unstabilized weighted model produced estimates with slightly wider confidence intervals implying large standard errors due to the high variability in the weights, unlike the stabilized weighted model.

3.8 Summary

This chapter has reviewed the Marginal Structural Model for repeated measured outcomes. We have applied an MSM to estimate the causal effect of HIV positivity awareness on condom use and multiple sexual partners and on marital dissolution using the Malawi Longitudinal Study of Families and Health (MLSFH). The findings have shown that HIV positivity awareness resulted an increased condom use. The results have further shown that awareness of HIV positivity led to an increase in marriage dissolution. We also found that among women, awareness of HIV positivity was associated with multiple sexual partners. The study found that both the standard binary logistic regression and the MSM showed an increase in condom use among individuals aware of their HIV positivity. Some studies have also shown similar effects between the standard regression model and the MSM (Suarez, Borràs, & Basagaña, 2011; Lusivika-Nzinga, Selinger-Leneman, Grabar, Costagliola, & Carrat, 2017; Yang, Eaton, McAlindon, & Lapane, 2015; Pega, Blakely, Glymour, Carter, & Kawachi, 2016). This may indicate less influence of prior exposure and time-varying confounders on the association between HIV positivity awareness and condom use. In our study, this could be attributed to education in years not being an important time-varying confounder for the association. However, using the standard multivariable logistic model, awareness of HIV positivity was associated with increased multiple sexual partners, a finding only found among women when using the MSM approach. We believe the differences in the finding could be due to the use of the inverse probability treatment weights and correct adjustment for time-dependent confounders (age and education level) (Hernan et al., 2002; VanderWeele, Jackson, & Li, 2016; Tsai et al., 2010).

Apart from an outcome being collected from an individual repeatedly over time, multiple outcomes at a single time-point are widely collected in biomedical science. Therefore, jointly estimating the effects on these outcomes would be of interest. In the next chapter, we describe statistical methods that can be used to assess treatment effectiveness on multiple outcomes.

CHAPTER 4 : Estimating Treatment Effects on Multiple Outcomes

4.1 Introduction

In the previous chapter, we reviewed causal inference methods for univariate longitudinal outcomes. Often times in health research one collects multiple outcomes at a single time point. Similar to methods in the previous chapter, jointly estimating the effects for the multiple outcomes would require accounting for the dependency between the outcomes. This chapter compares the performance of four multivariate regression methods, namely the multivariate outcome distribution, marginal mean, shared random effect and latent variable models through a simulation study. Their applicability is illustrated with a real-life example from a child growth study where the child's height-for-age z-score, weight-for-height z-score, and body-mass-index-for-age z-score could be correlated. A paper under review: Twabi, H. S., Manda, S. O., Small, D., & Kohler, H. (2023a), "A comparative simulation study of multivariate regression models for multiple outcomes" is based on material from this chapter

4.2 Structure of multivariate regression methods

In this section, we describe the regression structure that have been used to estimate the effect of treatment exposure and predictors on multiple outcomes.

4.2.1 Regression for univariate outcomes

Suppose for subject i, $i = 1, \ldots, n$ we observe an $(n \times p)$ covariate matrix denoted as $\mathbf{X}_i$ and a continuous outcome Y_i. A multiple regression model for the single outcome Y as a function of the p predictors for the n subjects is written as;

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \ldots & \vdots & \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

In vector notations this can be written as

$$\mathbf{Y}_i = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}_i \tag{4.1}$$

where $\boldsymbol{\epsilon}_i \sim N(0, \sigma^2)$

Estimation of the parameters in $\boldsymbol{\beta}$ is commonly done using the least squares approach. However, the maximum likelihood approach is also used.

4.2.2 Regression for Multivariate outcomes

Now suppose there J outcomes, such that subject i has $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{iJ})$ vector of outcomes. Assume that Y_{ij} are continuous and follow a multivariate normal distribution $(\mathbf{Y}_{ij} \sim N_j(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_j))$ where $\boldsymbol{\Sigma}_j$ is the variance-covariance matrix. The regression equation

for the multivariate outcomes will be given us;

$$\begin{pmatrix} y_{11} & y_{12} & \dots y_{1J} \\ y_{21} & y_{22} & \dots y_{2J} \\ \vdots & \vdots & \ddots \quad \vdots \\ y_{n1} & y_{n2} & \dots y_{nJ} \end{pmatrix} = \begin{pmatrix} 1 & x_{111} & x_{121} & \dots & x_{1p1} \\ 1 & x_{211} & x_{221} & \dots & x_{2p1} \\ & \vdots & \vdots & \vdots & \ddots \quad \vdots \\ 1 & x_{n11} & x_{n21} & \dots & x_{np1} \end{pmatrix} \begin{pmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0J} \\ \beta_{11} & \beta_{12} & \dots & \beta_{1J} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p1} & \beta_{p2} & \dots & \beta_{pJ} \end{pmatrix}$$

$$+ \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \dots & \epsilon_{1J} \\ \epsilon_{21} & \epsilon_{22} & \dots & \epsilon_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \dots & \epsilon_{nJ} \end{pmatrix}$$

This can be simplified and written in vector notation as

$$\begin{bmatrix} \mathbf{Y}_{1j} \\ \mathbf{Y}_{2j} \\ \vdots \\ \mathbf{Y}_{nj} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{11j} & \mathbf{x}_{12j} & \dots & \mathbf{x}_{1pj} \\ \mathbf{x}_{21j} & \mathbf{x}_{22j} & \dots & \mathbf{x}_{2pj} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{x}_{n1j} & \mathbf{x}_{n2j} & \dots & \mathbf{x}_{npj} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_{0j} \\ \boldsymbol{\beta}_{1j} \\ \vdots \\ \boldsymbol{\beta}_{pj} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_{1j} \\ \boldsymbol{\epsilon}_{2j} \\ \vdots \\ \boldsymbol{\epsilon}_{nj} \end{bmatrix} \tag{4.2}$$

This can be further simplified as

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \tag{4.3}$$

Estimation of the parameters in $\mathbf{B}$ can be done using several estimation approaches such as the ordinary least squares estimation, the maximum likelihood technique, methods of moments and bayesian estimation approaches (Anderson & Olkin, 1985; Arvanitoyannis, Katsota, Psarra, Soufleros, & Kallithraka, 1999; Yu & Wang, 2021).

4.3 Statistical regression models for multivariate outcomes

There are various statistical approaches that can be used to analyse multiple outcomes. In this section, we briefly describe relevant statistical methods that have been proposed to analyse multiple correlated outcomes.

4.3.1 Separate analysis

A simple approach used in analyzing multiple outcomes is to analyze each outcome separately by regressing each outcome on the predictors. The predictors included in the model also act as covariate adjustment to control for possible confounding. The regression models depend on the type of the outcome that is being modeled (Teixeira-Pinto & Mauri, 2011). For example, for continuous outcomes, a linear regression model is typically assumed, while for binary outcomes, a logistic or probit regression models is used. One limitation of separate regression analysis for multiple outcomes models is that it assumes no dependency between the outcomes and may provide biased effects (Verbeke, Fieuws, Molenberghs, & Davidian, 2014).

4.3.2 Multivariate normal model

The multivariate normal model provides an alternative to the separate regression models and accounts for the complex correlation structure between outcomes. The model assumes that the outcomes follow a normal distribution and are modelled as a multivariate normal distribution. The full joint distribution of the outcomes is specified by their means, variances and covariances. The covariance structure captures the strength of correlation

between the outcomes. For a multivariate continuous outcome $\mathbf{Y}_i$, we assume that $\mathbf{Y}_i$ conditional on $\mathbf{X}_i$ follows a multivariate normal distribution, $(\mathbf{Y}_{ij}|\mathbf{X}_i) \sim N(\mathbf{X}^T\mathbf{B}, \mathbf{\Sigma})$. Thus, the probability distribution function of $\mathbf{Y}_i$ is

$$f(\mathbf{Y}_i|\mathbf{X}_i, \mathbf{B}, \mathbf{\Sigma}) = (2\pi)^(- \frac{J}{2})|\mathbf{\Sigma}|^-\frac{1}{2}\text{exp}(-\frac{1}{2}(\mathbf{Y}_i - \mathbf{X}^T\mathbf{B})^T\mathbf{\Sigma}^{-1}(\mathbf{Y}_i - \mathbf{X}^T\mathbf{B})) \quad (4.4)$$

where $\mathbf{Y}_i$ and $\mathbf{X}_i$ denote the i^{th} rows of $\mathbf{Y}$ and $\mathbf{X}$, respectively.

Estimation of $\mathbf{B}$ is based on the maximum likelihood on the full joint distribution. The log-likelihood given $(\mathbf{Y}, \mathbf{B}, \mathbf{\Sigma})$ is given as:

$$In\{L(\mathbf{B}, \mathbf{X}, \mathbf{\Sigma})\} = -\frac{1}{2}\sum_{i=1}^{n}(\mathbf{Y}_i - \mathbf{X}_i\mathbf{B}^T)^T\mathbf{\Sigma}^{-1}(\mathbf{Y}_i - \mathbf{X}_i\mathbf{B}^T) + c \quad (4.5)$$

where c is a constant and does not depend on $\mathbf{B}$

4.3.3 Multivariate Bernoulli model

One of the limitations of the multivariate normal model is that it assumes that the outcomes are normally distributed and jointly follow a multivariate normal distribution. Thus, for binary outcomes, the mulltivariate normal model would produce biased estimates. For binary outcomes that are correlated, the multivariate Bernoulli model is used. This model assumes that the outcomes follow a Bernoulli distribution with a probability distribution that takes only values of 0 and 1. Suppose Y_{ij} be a binary random variable taking values $\{0, 1\}$ for the j^{th} outcome $(j = 1, 2, \ldots, J)$ of the i^{th} $(i = 1, \ldots, n)$ subject. $Y_{ij} \sim Bern(p_{ij})$ Then $\mathbf{Y} = Y_1, Y_2, \ldots Y_J$ is a multivariate Bernoulli R.V which takes values $\{0, 1\}^J = \{0, 1\} \times \{0, 1\} \times \ldots \times \{0, 1\}$. Then $P(\mathbf{Y} = y) = p(y_1, y_2, \ldots, y_J) =$

$p_{0,0,\ldots,0}^{\prod_{j=1}^{J}(1-y_j)} \times p_{1,0,\ldots,0}^{y_1\prod_{j=2}^{J}(1-y_j)} \times p_{0,1,\ldots,0}^{y_1(1-y_1)\prod_{j=3}^{J}(1-y_j)} \times \ldots \times p_{1,1,\ldots,1}^{\prod_{j=1}^{J}y_j}$. Suppose we have three multiple binary outcomes that follow a Bernoulli distribution, then the joint probability distribution function (pdf) will be

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = P_{000}^{(1-y_1)(1-y_2)(1-y_3)} P_{100}^{y_1(1-y_2)(1-y_3)} P_{010}^{(1-y_1)y_2(1-y_3)} P_{001}^{(1-y_1)(1-y_2)y_3}$$
$$P_{110}^{y_1y_2(1-y_3)} P_{101}^{y_1(1-y_2)y_3} P_{011}^{(1-y_1)y_2y_3} P_{111}^{y_1y_2y_3}$$

To estimate the parameters of the multivariate Bernoulli model, maximum likelihood estimation is often used. The likelihood function is computed based on the observed data and the assumed multivariate Bernoulli distribution. The estimates of the probabilities are obtained by maximizing the likelihood function. Taking the log of the joint pdf and solving for the probabilities and then taking the exponential, the joint pdf is given as

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) =$$
$$\exp\Bigg(y_1 \ln\frac{P_{100}}{P_{000}} + y_2 \ln\frac{P_{010}}{P_{000}} + y_3 \ln\frac{P_{001}}{P_{000}} + y_1y_2 \ln\frac{P_{110}P_{000}}{P_{100}P_{010}} +$$
$$y_1y_3 \ln\frac{P_{101}P_{000}}{P_{100}P_{001}} + y_2y_3 \ln\frac{P_{011}P_{000}}{P_{010}P_{001}} + y_1y_2y_3 \ln\frac{P_{010}P_{001}P_{111}}{P_{110}P_{101}P_{011}} + \ln P_{000} \Bigg) \tag{4.6}$$

for n subjects, the log-likelihood is written as

$$l(y) = \sum_{i=1}^{n} \Bigg(y_1 \ln \frac{P_{100}}{P_{000}} + y_2 \ln \frac{P_{010}}{P_{000}} + y_3 \ln \frac{P_{001}}{P_{000}} + y_1 y_2 \ln \frac{P_{110} P_{000}}{P_{100} P_{010}} + y_1 y_3 \ln \frac{P_{101} P_{000}}{P_{100} P_{001}} + y_2 y_3 \ln \frac{P_{011} P_{000}}{P_{010} P_{001}} + y_1 y_2 y_3 \ln \frac{P_{010} P_{001} P_{111}}{P_{110} P_{101} P_{011}} + \ln P_{000} \Bigg) \tag{4.7}$$

For expression 4.7 the link functions $\text{In}\frac{P_{000}P_{110}}{P_{100}P_{010}}$, $\text{In}\frac{P_{101}P_{000}}{P_{100}P_{001}}$, $\text{In}\frac{P_{000}P_{011}}{P_{010}P_{001}}$ and $\text{In}\frac{P_{010}P_{001}P_{111}}{P_{110}P_{101}P_{011}}$ model the dependency between the outcomes.

Some studies have used the odds ratio to link the dependency link function to predictors of interest. Other studies have used the marginal model and conditional model to estimate the effect of a predictor on the dependent outcomes. Dependency between the dependent binary outcomes can be measured using marginal probabilities or by specifying the correlation. (Lipsitz, Laird, & Harrington, 1991), (Liang, Zeger, & Qaqish, 1992) and (Carey, Zeger, & Diggle, 1993) proposed the use of a marginal odds ratio instead of specifying the marginal correlations.

4.3.4 Shared Random Effect model

Rather than specifying a multivariate distribution on the outcome and covariance vectors, a generalised linear mixed-model (GLMM) approach can be used. One example of a GLMM model is the shared random effect model. In this model, an unobserved random effect is specified to capture the association between the outcomes. The a random effect is assumed to be shared among the outcomes, and captures the unobserved variability common to all the outcomes. The random effects are usually assumed to follow a normal distribution.

The shared random effect model for J outcomes can be defined as;

$$g(\boldsymbol{\mu}_{ij}) = \beta \mathbf{X}_i + \mathbf{Z}_i b_i + \boldsymbol{\epsilon}_{ij} \tag{4.8}$$

where $g(.)$ is a known link function, $\mathbf{X}_i$ is a $n \times p$ and $\mathbf{Z}_i$ is a $n \times q$ vectors of covariates of interest. b_i is the shared random effect which captures the correlation between outcomes for subject i. b_i is distributed as $N(0, \mathbf{\Sigma_b})$, where $\Sigma_b = \sigma_b^2 \mathbf{I}$ and ϵ_{ij} is the error term and is distributed as $N(0, \mathbf{\Lambda}_i)$, where $\mathbf{\Lambda}_i = \sigma_e^2 \mathbf{I}$ and $\mathbf{I}$ is a $J \times J$ identity matrix. The error terms and the random effects are assumed to be independent $\text{Cov}(\epsilon_{ij}, b_i) = 0$. The covariates, $\mathbf{X}_i$, in equation 4.8 are assumed to be the same for the J outcomes. For continuous outcomes, the identity link is used ($g(E[Y_{ij}|X_i, b_i]) = E[Y_{ij}|X_i, b_i] = \boldsymbol{\beta} X_i + \mathbf{Z}_i b_i$) and for binary outcomes, the logit ($\Pr(Y_{ij}|\mathbf{X}_i, b_i) = \text{logit}^{-1}(\beta_{j1}\mathbf{X}_i + \mathbf{Z}_i b_i)$) or probit link is used. Conditional on the random effects $\mathbf{Y}_i|b_i \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_i b_i, \mathbf{\Lambda}_i)$, where a common variance is assumed for ϵ_{ij}. The marginal distribution of $\mathbf{Y}_i$ can be obtained by integrating out the unobserved random effect as follows:

$$\begin{aligned} f(Y_{ij}) &= \int f(Y_{ij}, b_i) db_i \\ &= \int \prod_{j=1}^{J} f(Y_{ij}|b_i) f(b_i) db_i \end{aligned}$$

The marginal distribution of $\mathbf{Y}$ is therefore distributed as, $\mathbf{Y}_i \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{Z}_i^T \Sigma_b \mathbf{Z}_i + \mathbf{\Lambda}_i)$

The parameters for equation 4.8 can be estimated using the maximum likelihood, restricted maximum likelihood or Bayesian estimation. Maximization can be done on the full joint likelihood or on the marginal likelihood. The marginal likelihood is computationally better

as the unobserved random effects are integrated out.

If we define $\boldsymbol{\theta}$ to be a vector of all the parameters in the model 4.8, $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)$, where $\boldsymbol{\alpha}$ is a vector of parameters used to calculate $\mathbf{V}_i$, where $\mathbf{V}_i = \mathbf{Z}_i^T \boldsymbol{\Sigma}_b \mathbf{Z}_i + \boldsymbol{\Lambda}_i$. Maximization of the marginal likelihood estimation is achieved through ;

$$L(\boldsymbol{\theta}) = \int \prod_{i=1}^{n} \prod_{j=1}^{J} (2\pi)^{-\frac{J}{2}} |\mathbf{V}_i|^{-\frac{1}{2}} \exp(-\frac{1}{2}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})^T \mathbf{V}_i^{-1}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})) \times \exp(-\frac{1}{2}\mathbf{b}_i \Lambda_i^{-1} \mathbf{b}_i^T) \tag{4.9}$$

Maximization is usually done on the log-likelihood

$$l(\boldsymbol{\theta}) = \int \sum_{i=1}^{n} \sum_{j=1}^{J} \left\{ -\frac{J}{2}\log(2\pi) - \frac{1}{2}\log(|\mathbf{V}_i|) - \frac{1}{2}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta})^T \mathbf{V}_i^{-1}(\mathbf{Y}_i - \mathbf{X}_i\boldsymbol{\beta}) - \frac{1}{2}\mathbf{b}_i \Lambda_i^{-1} \mathbf{b}_i^T \right\} \tag{4.10}$$

For binary outcomes $\mathbf{Y}_{ij}$ with values, 0 or 1, the likelihood function for $Pr(Y_{ij} = 1|\mathbf{X}_i, b_i, \boldsymbol{\beta})$ will be written as

$$\begin{aligned} L(x) &= \prod_{i=1}^{n} Pr(Y_{ij} = 1|\mathbf{X}_i, b_i, \boldsymbol{\beta}) \\ &= \prod_{i=1}^{n} \left\{ \int Pr(Y_{ij} = 1|\mathbf{X}_i, b_i, \boldsymbol{\beta}) f(b_i) \right\} \\ &= \prod_{i=1}^{n} \left\{ \int \prod_{j=1}^{J} Pr(Y_{ij} = 1|\mathbf{X}_i, b_i, \boldsymbol{\beta}) f(b_i) \right\} \\ &= \prod_{i=1}^{n} \left\{ \int \prod_{j=1}^{J} \frac{e^{\beta_j X_{ij} + b_i}}{1 + e^{\beta_j X_{ij} + b_i}} f(b_i) \right\} \end{aligned} \tag{4.11}$$

(4.12)

To estimate the β_j's and σ_b^2 we maximize the log of the likelihood in equation 4.12. However, the integrals of equation 4.12 do not have a closed-form solution and hence maximization of log-likelihood is done using approximation. One approximation approach is the Gauss-

Hermitian adaptive quadrature approach. Several softwares have the capacity of solving solutions for such equations. In R, the `lme4` package (Bates, Mächler, Bolker, & Walker, 2015) uses the adaptive Gaussian quadrature (AGQ) to approximate the integral of the likelihood. The default uses one quadrature point (nAGQ=1) which corresponds to the Laplace approximation method.

4.3.5 Marginal mean model

The shared random effect model provides estimates that are conditional and are interpreted as subject-specific. One may be interested in estimating the effect of an exposure on the average population, hence, a marginal mean model provides tat alternative. The marginal mean model approach models the mean of the multiple outcomes (at a population-level) as a function of the exposure $\mathbf{X}_i$ as follows,

$$\mu_{ij} = E(Y_{ij}) = g^{-1}(\mathbf{X}_i^T \boldsymbol{\beta}_j) \tag{4.13}$$

where μ_{ij} is linked to the exposure through the link function $g(.)$. where $\mathbf{X}_i$ is an $n \times p$ matrix of treatment exposures and predictors, $\boldsymbol{\beta}_j$ is a $p \times 1$ vector of coefficients relating $\mathbf{X}_i$ to Y_{ij}.

Estimation of the parameters for model 4.13 can be done either using a generalised least squares (GLS), maximum likelihood estimation (MLE) on the full joint distribution, Bayesian estimation approaches and generalised estimating equations. The GLS and the MLE perform efficiently when all outcomes follow a normal distribution, which may not be the case in real-life applications. The Bayesian estimation is more effective however,

require prior information. Specification of a full joint distribution on the mean of the multiple outcomes may require complex computation and may cause convergence problems when dealing with outcomes that are non-linear.

Since specification of the multivariate distribution of the outcomes can be complex, semi-parametric approaches such as the generalized estimating equations (GEE) is commonly used (Liang & Zeger, 1986; Lipsitz et al., 1991) for marginal mean models. Using the GEE approach, we only specify the functional relationship between the mean outcome and the exposure, and the mean and variance of the outcomes.

Using the GEE, the association between the outcomes within a subject are accounted for by defining a working correlation matrix $\mathbf{R}_i(\alpha)$ (Zeger, Liang, & Albert, 1988). where α is the specified correlation parameter to be estimated. The variance of Y_{ij} is modeled as $Var(Y_{ij}) = \phi\nu$, where ν is the variance function that depends on the mean and ϕ is the scale parameter. For example for continuous outcomes the variance of the multivariate outcomes, $Var\mathbf{Y}_i = \mathbf{V}_i$. $\mathbf{V}_i$ is the variance-covariance matrix and equals $\mathbf{A}_i^{\frac{1}{2}}\mathbf{R}_i(\alpha)\mathbf{A}_i^{\frac{1}{2}}$ and $\psi = 1$. $\mathbf{A}_i$ is the $J \times J$ diagonal matrix of the variance, where for a normally distributed continuous $\mathbf{Y}$ we have diag(Var$(Y_{ij})) = \sigma^2$. For binary outcomes that follow a Bernoulli distribution, $\mu_{ij} = P(Y_{ij} = 1)$, and $0 < \mu_{ij} < 1$ and the common link function is the logit link, where logit$\mu_{ij} = \log(\mu_{ij}/(1 - \mu_{ij})) = \mathbf{X}_i^T\boldsymbol{\beta}_j$. Therefore, $\mu_{ij} = \frac{\exp(\mathbf{x}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i^T\boldsymbol{\beta})}$ and $Var(Y_{ij}) = \mu_{ij}(1 - \mu_{ij})$. The coefficient parameter vector $\boldsymbol{\beta}$ for the marginal model can be obtained by solving the p estimating equations (Liang & Zeger, 1986; Zeger & Liang,

1986)

$$U_p(\boldsymbol{\beta}) = \sum_{i=1}^{N} \mathbf{D}_i^T \mathbf{V}_i^{-1}(\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0} \tag{4.14}$$

where $D_i = \frac{\partial(\boldsymbol{\mu}_i)}{\partial \boldsymbol{\beta}}$ and $\mathbf{V}_i = \mathbf{A}_i^{\frac{1}{2}} \mathbf{R}_i(\alpha) \mathbf{A}_i^{\frac{1}{2}}$. $\mathbf{R}_i(\alpha)$ is fully specified by the parameter α. The model is fit with the assumption that the correlation matrix is also known as a working correlation matrix. The coefficient parameters can be estimated based on the pairwise correlation between the outcomes to improve efficiency of the standard errors (Liang et al., 1992). There is no determined way of choosing a specific working correlation matrix and is left to the discretion of the researcher, this is one of the drawbacks of the generalized estimating equation (GEE) (Swan, 2006). The standard choices of the working correlation matrix are independent, unstructured, exchangeable and first order auto-regressive (AR(1)). For the independence $\mathbf{R}_i(\alpha)$, the pairwise correlation between two outcomes $corr(y_{ij}, y_{ik}) = 0, j \neq k$ and R_i is equal to the identity matrix. For the exchangeable correlation assumption, we assume that the correlation between the outcomes is the same $corr(y_{ij}, y_{ik}) = \alpha$. For the unstructured correlation assumption, the correlation between the outcomes has no specific structure, the pairwise association is different, $corr(y_{ij}, y_{ik}) = \alpha_{jk} = \alpha_{kj} = corr(y_{ik}, y_{ij})$. While the auto-regressive one assumes that the pairwise correlation increases if the outcomes are closer in time and decreases as the distance getting farther between time points, $corr(y_{ij}, y_{ik}) = \alpha^{|j-k|}$, which is mostly applied for repeated measures data.

The common approach of finding a solution to the generalized estimating equation is using the generalized weighted least squares or by solving the equations iteratively using the Fishers scoring method or Newton-Raphson iteration until convergence. One of the

advantages of the GEE is the parameter estimates in GEE model are consistent even when the association structure is misspecified. We used the Generalized estimating equations (GEE) proposed by Liang and Zeger (Liang & Zeger, 1986) assuming an exchangeable working correlation matrix and using an iterative process to estimate the parameter vector β.

4.3.6 Latent variable (LV) Model

Similar to the SRE, the latent variable model is a GLMM which is used to capture the unobserved effects that result in the dependency between the outcomes. The LV model assumes that the outcomes are correlated due to the unobserved latent variable, and estimates the effect of the treatment exposure and predictors on the latent variable, which is then used to estimate the effect on the individual outcomes (Sammel, Ryan, & Legler, 1997). The J outcomes can be modeled as a function of a single latent factor, a vector of fixed covariates $\mathbf{X}_i$ and the random error ϵ_{ij} as follows:

$$Y_{ij} = \mathbf{X}_{ij}\beta_j + \boldsymbol{\eta} u_i + \epsilon_{ij} \tag{4.15}$$

where β_j is the regression coefficient measuring the effect of the exposure on the specific outcome, $\boldsymbol{\eta}$ is a vector of $J \times 1$ factor loadings, $\epsilon_{ij} \sim N(0, \Sigma_\epsilon)$ and $u_i \sim N(0, \Sigma_u)$. u_i and ϵ_{ij} are independent implying $\text{Cov}(u_i, e_{ij}) = 0$. Model 4.15 assumes that the multiple outcomes are independent given the latent variables, which means that the dependency between the outcomes is due to the shared latent variable u_i.

The latent variables in model 4.15 reflect a subject's unobservable factors and are linked

to observed exposures unlike in the random effects model. Thus, we model latent factor u_i as a linear function of the predictor of interest (exposure) T_i:

$$u_i = T_i\alpha + \delta_i \tag{4.16}$$

where $\delta_i \sim N(0, \mathbf{I})$ and α measures the effect of the observed predictors (exposure) on latent variables and then on outcomes (Zhou, Lin, Song, & Li, 2014). Therefore, the model can be written as follows;

$$Y_{ij} = X_i\beta + \boldsymbol{\eta}T_i\alpha + \boldsymbol{\eta}\delta_i + \epsilon_{ij} \tag{4.17}$$

Therefore, the mean of Y_{ij} is given as $E[Y_{ij}] = X_i\beta + \boldsymbol{\eta}\alpha T_i$ and variance of Y_{ij} is Var$(Y_{ij}) = \boldsymbol{\eta}^T\boldsymbol{\eta} + \Sigma_\epsilon$. where diag$\{\Sigma_{epsilon}\} = \sigma^2$.

We could look at the joint distribution of the outcomes and the latent variable

$$f(Y_{ij}, u_i) \sim N\left(\begin{pmatrix} X_i\beta + \eta\alpha T_i \\ T_i\alpha \end{pmatrix}, \begin{pmatrix} \boldsymbol{\eta}^T\boldsymbol{\eta} + \Sigma_{epsilon} & \boldsymbol{\eta} \\ \boldsymbol{\eta}^T & 1 \end{pmatrix}\right)$$

The parameters $\eta, \alpha, \beta, \Sigma_{epsilon}$ are estimated using the maximum likelihood (ML). The complete joint likelihood for an individual is given as

$$L(\alpha, \eta, \beta, \Psi) = \prod_{i=1}^{n}\prod_{j=1}^{J} f(y_{ij}|\alpha, \beta, \Psi, \eta) f(u_i, \alpha)$$

$$= \prod_{i=1}^{n}\prod_{j=1}^{J} \frac{1}{(2\pi)^n}|\Sigma_{epsilon}|^{-1}\text{exp}(-\frac{1}{2}(y_{ij} - X_i\beta - \alpha T_i\eta)^T\Sigma_{epsilon}^{-1}(y_{ij} - X_i\beta - \alpha T_i\eta))$$

$$\times\frac{1}{(2\pi)} \times \Sigma_u \times -\text{exp}(\frac{1}{2}(u_i - T_i\alpha)^T(u_i - T_i\alpha)) \tag{4.18}$$

Therefore, the log-likelihood is given as;

$$\log L(\alpha, \eta, \beta, \Sigma_{epsilon}) = \sum_{i=1}^{n}\sum_{j=1}^{J} \frac{1}{(2\pi)^n} + \log(|\Sigma_{epsilon}|^{-1}) - \frac{1}{2}(y_{ij} - X_i\beta - \alpha T_i\eta)^T\Sigma_{epsilon}^{-1} \quad (4.19)$$
$$(y_{ij} - X_i\beta - \alpha T_i\eta) + \frac{1}{(2\pi)} + \ln|\mathbf{\Sigma_u}| - \frac{1}{2}(u_i - T_i\alpha)^T(u_i - T_i\alpha)$$

The maximum likelihood can be used to estimate the parameters of model 4.17 if the log of the likelihood has a closed-form solution, particularly if the outcomes are distributed normally, and the latent variable is assumed to follow a normal distribution. Another approach would be to use the Expectation-Maximization (EM) algorithm (Sammel & Ryan, 1996; Sammel et al., 1997). The restricted maximum likelihood has also been used to estimate the variance of the parameters, as using ML, the estimated variance is biased. We apply an approach by (Teixeira-Pinto & Mauri, 2011) using the maximum likelihood with a closed form for the outcomes that follow a normal distribution. However, for the binary outcomes, the latent variable model is written as

$$\text{logit}\{P(Y_{ij} = 1|, \alpha, \boldsymbol{\beta}, u_i)\} = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\eta}T_i\alpha + \mathbf{\Lambda}\delta_i \quad (4.20)$$

The likelihood will be defined as

$$L(Y_{ij}) = \prod_{i=1}^{n} f(Y_{ij}|\beta, u_i, \alpha)f(u_i, \alpha)du_i \quad (4.21)$$

where $f(Y_{ij}|\beta, \alpha, \Lambda) = \prod_{j=1}^{J} p_{ij}^{Y_{ij}}(1 - p_{ij}^{Y_{ij}})$ and $p_{ij} = P(Y_{ij} = 1|, \alpha, \boldsymbol{\beta}, u_i)$. Therefore, the log of the likelihood can be written as;

$$= \prod_{i=1}^{n}\prod_{j=1}^{J}[P(Y_{ij} = 1|, \alpha, \boldsymbol{\beta}, u_i)]^{y_{ij}}[(1 - P(Y_{ij} = 1|, \alpha, \boldsymbol{\beta}, u_i)]^{1-y_{ij}} \times f(u_i) \quad (4.22)$$

Where the distribution of u_i can follow an assumed distribution, the common distribution assumed is that of the normal or Bernoulli. The integral for the log-likelihood in equation 4.22 does not have a closed-form solution. Hence, maximization of the log-likelihood is done using a numerical approximation such as the Gaussian adaptive quadrature to estimate maximum-likelihood parameters. The Stata package `gllamm` can be used to fit the latent variable model (Rabe-Hesketh, Skrondal, & Pickles, 2004). Since u_i is unobservable, the estimate of u_i can be obtained by using the posterior mean $E[u_i|Y_j]$ through expectation-maximization technique (Sammel & Ryan, 1996; Sammel et al., 1997).

4.4 Simulation Study

We performed a simulation study to assess the performance of the univariate, multivariate, marginal, shared random effect and latent variable models in estimating the effect of an exposure on the multiple outcomes by assessing the bias and 95% Coverage for the estimated effects. Several scenarios were considered by varying the outcome type (continuous or binary or mixed), the number of outcomes ($J = 2, 3$), the correlation between outcomes (no correlation $\rho = 0$, low $\rho = 0.3$, and high $\rho = 0.8$), the sample sizes (200, 500, 1000), and the covariate of interest (same or different). In each simulation we generated multiple outcomes (continuous or binary or mixed). The covariate X_1, was generated from Bernoulli, $X_1 \sim Bernoulli(0.5)$ and was considered the exposure of interest. For each simulation we generated 1000 independent data sets. The outcomes were simulated under three scenarios, as continuous, binary or a mix of both binary and continuous.

Continuous outcomes

We used the following model to simulate values for two continuous outcomes that follow a bivariate normal distribution $\mathbf{Y}_i = (Y_{i1}, Y_{i2})^T$

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i \tag{4.23}$$

For similar covariate effects across the outcomes, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)^T$, where $\boldsymbol{\beta}_1 = \begin{pmatrix} \beta_{01} \\ \beta_{11} \end{pmatrix}$ and β_{11} is the exposure effect for Y_1 and $\boldsymbol{\beta}_2 = \begin{pmatrix} \beta_{02} \\ \beta_{12} \end{pmatrix}$, β_{12} is the estimated effect for Y_2 . For similar exposure effects we define $\boldsymbol{\beta}_j = (\beta_{0j} = -1.1, \beta_{1j} = 0.5)^T$.

For two outcomes with different covariate effects $\beta_1 = (-1.1, 0.5)$ and $\beta_2 = (-0.2, 0.4)$.

Dependency between the outcomes was captured by assuming a joint distribution on ϵ_{ij}. ϵ_i are errors from a multivariate normal distribution.

$$\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2})^T \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

where the correlation matrix is obtained from the covariance matrix with $\rho = \sigma_{jl}/\sigma_j\sigma_l$ and σ_j and σ_l are equal to one. This approach was extended in a similar manner to simulate three continuous outcomes. For three continuous outcomes with different effects, we had $\beta_1 = (-1.1, 0.3)$, $\beta_2 = (-0.2, 1.5)$ and $\beta_3 = (-0.7, 1.0)$.

Binary outcomes

We simulated two binary outcomes $\mathbf{Y}_{ij} = (Y_{i1}, Y_{i2})$ using the bindata package in R (Leisch, Weingessel, & Hornik, 1998). The outcomes are firstly generated from a bivariate normal distribution and thereafter transformed into binary outcomes. We generate the binary outcomes (for Y_1, Y_2) such that the marginal probabilities satisfy

$$P(Y_{ij} = 1|\mathbf{X}_i, T) = [1 + \exp(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\mathbf{X}_1)]^{-1}, j = 1, 2$$

where for similar covariate effects we set $\boldsymbol{\beta} = (\beta_{oj} = 1.0, \beta_{1j} = -0.2)^T$ and for outcome-specific efects we set $\boldsymbol{\beta} = (\beta_{01} = 1.0, \beta_{11} = 0.1, \beta_{21} = 0.2, \beta_{31} = 0.3)^T$. Given the coefficients values, we obtained the marginal probabilities for Y_1 and Y_2, respectively. The correlated binary data were generated by specifying the marginal probabilities based on the given marginal probability and ρ. A similar approach was used to simulate three correlated binary outcomes with the following coefficients for different covariate effects ($\beta_{11} = 1.0, \beta_{12} = 0.5, \beta_{13} = 0.3$).

Mixed outcomes

We simulate two mixed outcomes where Y_1 was continuous and Y_2 was binary associated with a specific predictor effect. Data were generated from a bivariate normal distribution as follows

$$\begin{pmatrix} y_{1i} \\ y^*_{2i} \end{pmatrix} \sim MVN \left(\begin{pmatrix} -0.1 + 0.5X_{11} \\ -0.3 + 0.2X_{12} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

For the exposure effects we set $\boldsymbol{\beta} = (\beta_{11} = 0.5, \beta_{12} = 0.3)$. We created the binary variable y_{2i} from the normally distributed variable y_{2i}^* using the following relation:

$$y_{2i} = \begin{cases} 0 & \text{if } y_{2i}^* \leq 0 \\ 1 & \text{if } y_{2i}^* > 0 \end{cases}$$

A similar approach was used to simulate three correlated mixed outcomes (two continuous and one binary) with the following coefficients ($\beta_{11} = 0.5, \beta_{12} = 0.2, \beta_{13} = 0.4$).

4.4.1 Performance assessment and model fitting

Performance of the methods in estimating the effects was done by assessing the accuracy of the bias and 95% coverage under each scenario. The bias in the estimated effects was calculated as the difference between the average effect estimates across the 1000 replicates and the true effects.

All simulations and analyses were conducted in R (Dessau & Pipper, 2008). For the estimation of the effects using the marginal model, we used the `geepack` (Højsgaard et al., 2006). The `MVLM` and `mvProbit` was used for the multivariate normal and bernoulli models, respectively. The `JGEE` package (Inan, 2015) was used for estimating parameters for the marginal mean model with mixed outcomes. For generalized linear mixed models (GLMM), we used the `lme4` (Bates, Sarkar, Bates, & Matrix, 2007) and the Mcglm packages (Bonat, 2018). We used the GEE assuming an exchangeable working correlation matrix.

4.5 Simulation Results

4.5.1 Continuous Outcomes

4.5.1.1 Bias

For two outcomes with similar effects, the estimated effects for the univariate, multivariate, shared random effect and the marginal mean models were more accurate regardless of the assumed correlation strength and sample size compared to estimated effects for the latent variable model (Table 4.1, Figure 4.1). For three continuous outcomes with similar covariates effects, the estimated effects were more accurate for the univariate (for $\rho = 0$, $n > 200$), the marginal mean (for $rho \geq 0.3$ and $n > 200$) (Table 2 in the Appendix). For two or three continuous outcomes with different effects, the estimated effects were more accurate for all methods except that latent variable model (Tables 4.2, 4.3).

4.5.1.2 95% Coverage

For two or three continuous outcomes with similar effects, the 95% coverage were more than the nominal 95% for the marginal mean and shared random effect models for $\rho = 0.8$ and $n > 500$. We also observed 95% coverage more than the nominal 95% for $n = 500$ and $\rho = 0$. The estimated effects for the latent variable model were less accurate across all sample sizes regardless of the correlation strength and number of outcomes.

Table 4.1: Estimated effects, bias and 95% coverage for two continuous outcomes with same covariates. β_0=-1.1, β_1=0.5

Sample size	Model	$\widehat{\beta}_0$	$\widehat{\beta}_1$	Bias ($\widehat{\beta}_0$)	Bias ($\widehat{\beta}_1$)	95% Cov. ($\widehat{\beta}_0$)	95% Cov. ($\widehat{\beta}_1$)
$\rho = 0$							
	Univariate	-1.103	0.500	0.003	3.00E-04	96.3	93.7
	Multivariate	-1.096	0.496	-0.004	0.004	80.2	83.6
	Shared	-1.103	0.499	0.003	3.00E-04	96.1	93.1
n=200	Marginal	-1.103	0.499	0.003	3.00E-04	96.1	93.1
	Latent	-1.100	0.372	0.000	0.128	90.2	80.6
	Univariate	-1.101	0.504	0.001	-0.004	95.6	95.5
	Multivariate	-1.101	0.504	0.001	-0.004	95.6	95.5
	Shared	-1.103	0.504	0.001	-0.004	95	95.3
n=500	Marginal	-1.101	0.504	0.001	-0.004	95.3	95.4
	Latent	-1.099	0.375	0.001	0.126	73.9	66.6
	Univariate	-1.100	0.500	-4.00E-04	0.0012	94.4	94.4
	Multivariate	-1.099	0.498	-9.00E-04	0.002	19.5	17.3
	Shared	-1.099	0.499	-4.00E-04	0.0012	94.3	94.3
n=1000	Marginal	-1.099	0.499	-4.00E-04	0.0012	94.3	94.3
	Latent	-1.099	0.373	0.001	0.127	41	43.8
$\rho = 0.3$							
	Univariate	-1.095	0.497	-0.005	0.003	94.7	94.8
	Multivariate	-1.095	0.497	-0.005	0.003	18.3	18
	Shared	-1.099	0.497	-0.001	0.003	94.3	94.6
n=200	Marginal	-1.099	0.497	-0.001	0.003	94.3	94.6
	Latent	-1.097	0.330	-0.003	0.170	80.3	71.6
	Univariate	-1.097	0.497	-0.003	0.003	94.6	95
	Multivariate	-1.097	0.497	-0.003	0.003	19.6	20.5
	Shared	-1.099	0.501	-0.001	-0.001	94.4	94.8
n=500	Marginal	-1.099	0.501	-0.001	-0.001	94.4	94.8
	Latent	-1.098	0.330	-0.002	0.170	42.1	43.6
	Univariate	-1.099	0.498	0.00009	0.002	94.5	94.5
	Multivariate	-1.099	0.498	-9.00E-04	0.002	19.4	20.3
	Shared	-1.099	0.498	-0.001	0.002	94.5	94.3
n=1000	Marginal	-1.099	0.498	-0.001	0.002	94.5	94.3
	Latent	-1.099	0.331	0.001	0.169	8.7	17.4
$\rho = 0.8$							
	Univariate	-1.095	0.496	-0.005	0.004	94	93.9
	Multivariate	-1.098	0.496	-0.003	0.004	82.6	81.6
	Shared	-1.098	0.496	-0.003	0.004	93.7	93.7
n=200	Marginal	-1.098	0.496	-0.003	0.004	93.7	93.7
	Latent	-1.096	0.266	-0.004	0.234	40.4	40.3
	Univariate	-1.097	0.498	-0.003	0.002	95.9	95.4
	Multivariate	-1.098	0.500	-0.002	-2.00E-04	80.0	79.9
	Shared	-1.098	0.500	-0.002	-2.00E-04	95.7	95.2
n=500	Marginal	-1.098	0.500	-0.002	-2.00E-04	95.7	95.2
	Latent	-1.098	0.264	-0.002	0.236	1.9	4.4
	Univariate	-1.099	0.499	0.0007	0.001	94.4	94.4
	Multivariate	-1.099	0.499	-7.00E-04	0.001	82.8	83.1
	Shared	-1.099	0.498	-0.001	0.002	94.4	94.4
n=1000	Marginal	-1.099	0.498	-0.001	0.002	94.4	94.4
	Latent	-1.099	0.264	-0.001	0.236	0.0	0.0

Table 4.2: Estimates, Bias and 95% coverage for the estimated parameters of the two continuous outcomes for selected estimates $\widehat{\beta}_{11}$ and $\widehat{\beta}_{12}$ with different covariates, True values; $\beta_{11} = 0.5$, $\beta_{12} = 0.4$

		Estimates		**Bias**		**95% Coverage**	
sample size	Model	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$
$\rho = 0$							
	Univariate	0.4997	0.3961	0.0003	0.0039	96.3	93.7
	Multivariate	0.4997	0.3961	0.0003	0.0039	96.3	93.7
	Shared	0.500	0.396	-0.004	0.004	96	93.4
n=200	Marginal	0.500	0.396	-0.004	0.004	96.1	93.1
	Latent	0.373	0.269	-0.004	0.131	90.2	80.6
	Univariate	0.5037	0.4989	-0.0037	0.0011	95.6	80.3
	Multivariate	0.4997	0.3961	0.0003	0.0039	18.3	13.3
	Shared	0.373	0.269	-0.003	0.001	95.7	95.2
n=500	Marginal	0.504	0.399	-0.003	0.001	95.3	95.4
	Latent	0.371	0.266	-0.003	0.134	74.1	24.8
	Univariate	0.4988	0.4981	0.0012	0.0019	94.4	94.4
	Multivariate	0.4988	0.3981	0.0012	0.0019	19.5	13.5
	Shared	0.499	0.398	-0.001	0.002	94.4	93.8
n=1000	Marginal	0.499	0.398	-0.001	0.002	94.3	94.3
	Latent	0.367	0.267	-0.001	0.133	41.3	44.1
$\rho = 0.3$							
	Univariate	0.4967	0.497	0.0033	0.003	94.7	94.8
	Multivariate	0.496	0.397	0.0033	0.003	19.6	17.6
	Shared	0.497	0.396	-0.006	0.004	94.4	94.5
n=200	Marginal	0.497	0.396	-0.006	0.004	94.3	94.6
	Latent	0.289	0.188	-0.006	0.212	80.4	71.7
	Univariate	0.5013	0.497	-0.0013	0.003	94.6	95
	Multivariate	0.5013	0.397	-0.0013	0.003	19.4	16.8
	Shared	0.503	0.404	-0.001	-0.004	94.4	94.8
n=500	Marginal	0.503	0.404	-0.001	-0.004	94.4	94.8
	Latent	0.291	0.186	-0.003	0.214	42.5	43.6
	Univariate	0.4978	0.4992	0.0022	0.0008	94.5	94.5
	Multivariate	0.497	0.3992	0.0022	0.0008	18.3	15.5
	Shared	0.498	0.399	-0.001	0.001	94.5	94.5
n=1000	Marginal	0.498	0.399	-0.001	0.001	94.5	94.3
	Latent	0.286	0.187	-0.001	0.213	8.9	17.7
$\rho = 0.8$							
	Univariate	0.496	0.396	0.0038	0.0036	94	93.9
	Multivariate	0.496	0.396	0.0038	0.0036	17.4	14.5
	Shared	0.501	0.416	-0.005	0.004	93.0	93.3
n=200	Marginal	0.501	0.416	-0.005	0.004	93.7	93.7
	Latent	0.299	0.214	-0.005	0.206	44.1	44.2
	Univariate	0.500	0.398	0.0004	0.0022	95.9	80.5
	Multivariate	0.5002	0.3978	0.0002	0.0022	20	15.5
	Shared	0.501	0.397	-0.003	0.003	95.7	95.2
n=500	Marginal	0.501	0.397	-0.003	0.003	95.7	95.2
	Latent	0.297	0.193	-0.003	0.207	2.7	5.5
	Univariate	0.498	0.399	-1.598	0.1014	94.4	62
	Multivariate	0.498	0.399	0.0022	0.0014	17.2	12.9
	Shared	0.498	0.399	-0.001	0.001	94.4	94.4
n=1000	Marginal	0.498	0.399	-0.001	0.001	94.4	94.4
	Latent	0.293	0.194	-0.001	0.206	0.0	0.0

Table 4.3: Estimates, Bias and 95% coverage for three continuous outcomes with different covariates. True parameters β_{11}=0.3, β_{12}=1.5, β_{13}=1.0

		Estimates			**Bias**			**95% Coverage**		
sample size	Model	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{13}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{13}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{13}$
$\rho = \mathbf{0}$										
	Univariate	0.298	1.496	1.002	0.002	0.004	-0.002	94.8	95.7	94.3
	Multivariate	0.298	1.496	1.002	0.002	0.004	-0.002	11.3	43.1	30
	Shared R.E.	0.298	1.503	1.002	-0.001	-0.004	0.000	93.4	95.5	92.8
n=200	Marginal	0.298	1.503	1.002	0.001	0.004	0.000	94.4	95.3	94
	Latent	0.323	1.329	0.885	-0.024	0.173	0.118	9	11.5	10.4
	Univariate	0.309	1.497	0.995	-0.009	0.003	0.006	94.2	95.7	94.6
	Multivariate	0.309	1.497	0.9945	-0.009	0.003	0.005	11	45.5	35.4
	Shared R.E.	0.298	1.503	1.002	0.002	-0.003	-0.002	93.8	95.3	94.4
n=500	Marginal	0.298	1.503	1.002	-0.002	0.003	0.002	93.9	95.5	94.4
	Latent	0.323	1.329	0.885	-0.023	0.171	0.115	0.0	0.1	0.0
	Univariate	0.305	1.503	1.001	5.00E-04	-0.003	-0.001	94	96.2	95.7
	Multivariate	0.301	1.503	1.001	5.00E-0	-0.003	-0.001	11.3	47.2	33
	Shared R.E.	0.299	1.503	1.000	0.001	-0.003	0.000	94.1	96.1	95.1
n=1000	Marginal	0.299	1.503	1.000	0.001	-0.003	0.000	94	96.1	95.6
	Latent	0.325	1.329	0.884	-0.024	0.171	0.116	0.0	0.0	0.0
$\rho = \mathbf{0.3}$										
	Univariate	0.296	1.501	0.999	0.004	-0.001	-0.0008	93.8	95.8	93.2
	Multivariate	0.296	1.501	0.999	0.004	-0.001	0.0008	11.5	42.4	34
	Shared R.E.	0.300	1.505	1.002	0.000	-0.005	-0.002	93.4	95.5	92.8
n=200	Marginal	0.300	1.505	1.001	0.000	0.005	0.001	93.4	95.5	92.8
	Latent	0.332	1.258	0.836	-0.032	0.242	0.164	6.1	5.6	3.6
	Univariate	0.302	1.499	1.011	-0.002	-0.0005	-0.011	94	96.2	94.5
	Multivariate	0.302	1.4995	1.0106	-0.009	-0.004	-0.003	12	48	32.7
	Shared R.E.	0.297	1.504	1.003	0.000	-0.005	-0.002	94.0	96.2	94.2
n=500	Marginal	0.297	1.504	1.003	0.000	0.005	0.001	94	96.2	94.3
	Latent	0.331	1.259	0.839	-0.030	0.241	0.161	0.0	0.0	90.1
	Univariate	0.301	1.498	0.998	-0.001	0.002	0.002	96	95.4	94.5
	Multivariate	0.3014	1.4979	0.9984	-0.001	0.002	0.002	11.2	44.5	31.5
	Shared R.E.	0.298	1.504	1.001	0.002	-0.004	-0.001	96.0	95.4	94.5
n=1000	Marginal	0.298	1.504	1.001	-0.002	0.004	0.001	96	95.4	94.5
	Latent	0.332	1.259	0.838	-0.032	0.241	0.162	0.0	0.0	0.0
$\rho = \mathbf{0.8}$										
	Univariate	0.298	1.500	0.998	0.002	0.0002	0.003	93.6	94.8	94.2
	Multivariate	0.2979	1.5002	0.9975	0.002	0.0002	0.003	10.5	42.4	31.8
	Shared R.E.	0.297	1.506	1.006	0.003	-0.006	0.000	93.6	94.4	93.8
n=200	Marginal	0.297	1.506	1.005	0.003	-0.006	-0.005	93.6	94.4	93.8
	Latent	0.339	1.174	0.783	-0.039	0.327	0.218	1.4	0.7	0.2
	Univariate	0.309	1.504	1.003	-0.009	-0.004	-0.003	94.1	95.8	94.6
	Multivariate	0.3086	1.5044	1.0025	-0.009	-0.004	-0.003	12.1	44.4	36.2
	Shared R.E.	0.295	1.503	1.007	0.005	-0.003	-0.007	94.1	95.8	94.5
n=500	Marginal	0.295	1.503	1.007	0.005	-0.003	-0.007	94.1	95.8	94.5
	Latent	-0.038	0.338	1.173	0.785	0.327	0.216	0.0	0.0	0.0
	Univariate	0.299	1.498	0.999	0.0008	0.002	0.0001	94.8	95	95.1
	Multivariate	0.2992	1.4979	0.9999	0.0008	0.002	0.0001	9.4	44.3	34.4
	Shared R.E.	0.300	1.503	0.999	0.003	-0.003	-0.004	94.7	94.9	95.1
n=1000	Marginal	0.300	1.503	0.999	0.003	-0.003	-0.004	94.7	94.9	95.1
	Latent	0.340	1.173	0.783	-0.040	0.327	0.217	0.0	0.0	0.0

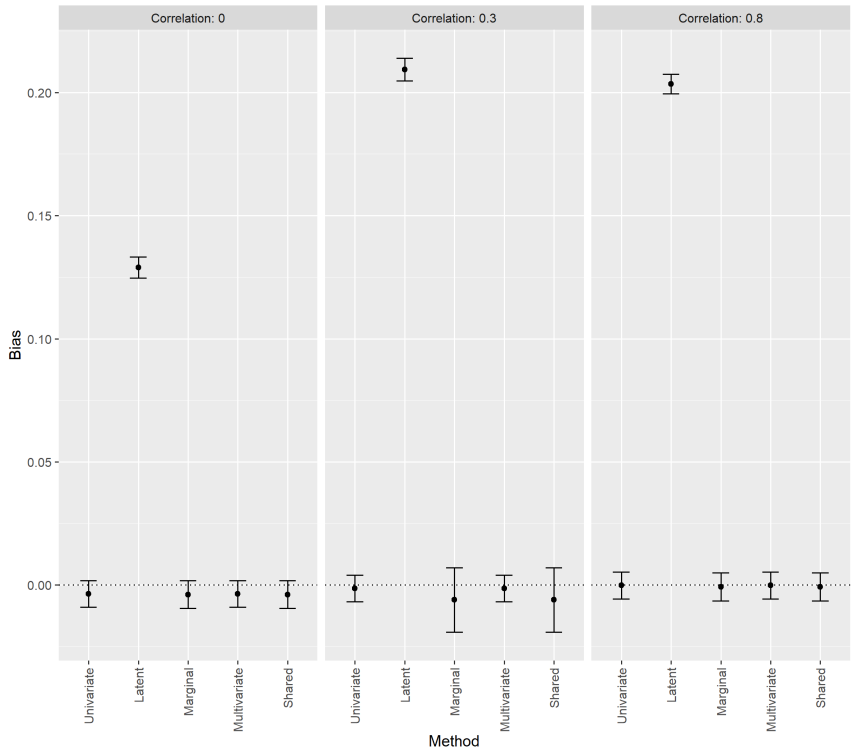


Figure 4.1: Mean treatment effect for two continuous outcomes with same effect,$\beta_{\text{continuous}} = 0.5$. **Source**: Researcher (2023)

4.5.2 Binary outcomes

4.5.2.1 Bias

For two binary outcomes with same covariate effects, the estimated effects were accurate for univariate and marginal mean models (across all correlation strengths) regardless of the sample size as shown in Tables 4.4 and 3 and Figure 4.2. However, the estimated effects for the shared RE were only accurate for $\rho = 0$, regardless of the sample size. As the assumed correlation increased ($\rho \geq 0.5$), the estimated effects for the shared RE were less accurate (Bias $>> 0$). The estimated effects for the latent variable and the multivariate Bernoulli model were less accurate for $\rho \geq 0$. For three outcomes with same effects the trivariate Bernoulli and marginal model were more accurate compared to the rest of the

models. For two binary outcomes with different effects, none of the models were accurate across the different sample sizes and correlation strength. While for three outcomes with different effects only the marginal mean model was accurate for $\rho \geq 0$, regardless of the sample size. The multivariate model was accurate for $n \geq 500$ and $\rho \leq 0.3$ (Tables 4.5, and 4.6).

4.5.2.2 95% Coverage

For two binary outcomes with similar effects, the coverage proportion of the estimated exposure effects was low for the all models except the marginal mean model. The marginal model had a coverage proportion closer to the nominal 95% for $\rho \geq 0.3$ and $n \geq 500$. There was an overcoverage for the marginal mean model for $\rho = 0.8$and$n = 1000$. For three outcomes with similar effects, the coverage proportion was closer to the nominal 95% for $\rho = 0, 0.3$ and $n > 200$ for the latent variable and marginal mean models. For two and three binary outcomes with different effects, only the marginal mean model had a coverage proportion closer to the nominal 95%, see Tables 4.5 and 4.6.

Table 4.4: Estimates, bias and 95% coverage for the effects on two binary outcomes with same effects. β_0=1.0, $\beta_1 = -0.2$

		Estimated parameters					
sample size	Model	$\widehat{\beta}_0$	$\widehat{\beta}_1$	Bias ($\widehat{\beta}_0$)	Bias ($\widehat{\beta}_1$)	95% Cov. ($\widehat{\beta}_0$)	95% Cov. ($\widehat{\beta}_1$)
$\rho = \mathbf{0}$							
	Univariate	1.004	-0.203	-0.0047	0.0029	92	93
	Multivariate	1.689	-0.713	-0.688	0.166	0.1	40.5
	Shared	1.040	-0.208	-0.040	8.00E-03	92.4	93.3
n=200	Marginal	1.017	-0.204	-0.017	4.00E-03	93	93.5
	Latent	1.699	-0.075	-0.699	-0.125	0.0	89.7
	Univariate	1.002	-0.198	-0.0016	-0.0017	89.5	91.3
	Multivariate	1.679	-0.359	-0.679	0.159	20.4	60.5
	Shared	1.027	-0.216	-0.027	0.016	90.0	91.4
n=500	Marginal	1.013	-0.213	-0.013	0.013	90.3	91.7
	Latent	1.672	-0.124	-0.672	-0.076	10.0	92.3
	Univariate	1.003	-0.203	-0.0034	0.0028	83.8	84.1
	Multivariate	1.671	-0.352	-0.671	0.152	1.5	40.7
	Shared	1.016	-0.203	-1.60E-02	0.003	31	89.2
n=1000	Marginal	1.002	-0.201	-2.00E-03	8.00E-04	84	84.3
	Latent	1.671	-0.156	-0.671	-0.044	0.0	88.6
$\rho = \mathbf{0.3}$							
	Univariate	1.003	-0.206	-0.003	0.0059	92.2	93
	Multivariate	1.683	-0.363	-0.683	0.163	90.4	88.2
	Shared	1.195	-0.226	-0.195	0.026	92.9	93.4
n=200	Marginal	1.010	-0.194	-0.011	-6.00E-03	94.5	93.9
	Latent	1.905	-0.100	-0.905	-0.100	0.5	94.9
	Univariate	1.001	-0.1977	0.0005	-0.0023	91.1	90.4
	Multivariate	1.681	-0.355	-0.681	0.155	85.3	88.9
	Shared	1.187	-0.023	-0.181	0.033	91.7	90.7
n=500	Marginal	1.006	-0.201	-0.006	6.00E-04	94.2	95.9
	Latent	1.840	-0.134	-0.840	-0.066	0.0	95.4
	Univariate	1.0044	-0.2025	-0.0044	0.0025	83.6	82.9
	Multivariate	1.673	-0.354	0.154	-0.674	89.2	88.4
	Shared	1.175	-0.229	-0.175	0.029	84.2	83.2
n=1000	Marginal	1.000	-0.199	-1.00E-04	-0.002	98	96.4
	Latent	1.826	-0.162	-0.826	-0.038	0.0	87.9
$\rho = \mathbf{0.8}$							
	Univariate	0.9979	-0.2049	0.0021	0.0049	91.7	92.1
	Multivariate	1.678	-0.358	-0.678	0.156	84.8	83
	Shared	1.686	-0.374	-0.686	0.174	86.4	90.3
n=200	Marginal	1.011	-0.196	-0.011	-0.004	93.2	94.5
	Latent	7.749	-1.088	-6.748	0.888	0.2	88.3
	Univariate	1.000	-0.1996	0.0002	0.0004	98.9	92.5
	Multivariate	1.676	-0.356	-0.676	0.156	85.4	84.4
	Shared	1.671	-0.352	-0.671	1.52E-01	90.3	91
n=500	Marginal	1.003	-0.195	-0.003	-5.00E-03	95.4	97.5
	Latent	7.749	-0.982	-6.749	0.782	0.0	80.2
	Univariate	1.001	-0.2011	-0.0029	0.0026	93.4	93.5
	Multivariate	1.675	-0.356	-0.675	0.156	82.0	89.2
	Shared	1.671	-0.357	-6.71E-01	0.157	84.2	84
n=1000	Marginal	1.000	-0.200	-4.00E-04	0.000	99	97.8
	Latent	7.776	-0.958	-6.776	0.758	0.0	93.8

Table 4.5: Estimates, bias and 95% Coverage for estimated parameters for two binary outcomes with different effects. True parameters $\beta_{11} = 0.1$, $\beta_{12} = 0.3$

		Estimates		**Bias**		**95% Coverage**	
Sample size	Model	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{12}$
$\rho = 0$							
	Univariate	0.169	0.484	-0.073	-0.205	94.7	90.4
	Multivariate	0.197	0.487	0.097	-0.187	93.4	90.1
	Shared	0.173	0.505	-0.073	-0.205	94.8	89.2
n=200	Marginal	0.169	0.484	-0.069	-0.183	95.2	90.8
	Latent	0.445	0.505	-0.345	-0.205	81.3	98.4
	Univariate	0.198	0.489	-0.124	-0.188	91.8	84.3
	Multivariate	0.174	0.490	0.074	-0.190	94.4	83.5
	Shared	0.201	0.500	-0.101	-0.200	90.3	82.0
n=500	Marginal	0.198	0.488	-0.098	-0.188	92.1	84.8
	Latent	0.463	0.501	-0.363	-0.201	58.5	95.7
	Univariate	0.187	0.488	-0.087	-0.188	92.2	70.4
	Multivariate	0.175	0.489	0.075	-0.189	93.4	84.4
	Shared	0.189	0.497	-0.089	-0.197	92.2	70.9
n=1000	Marginal	0.187	0.488	-0.087	-0.188	92.2	70.9
	Latent	0.451	0.502	-0.351	-0.202	31.3	93.1
$\rho = 0.3$							
	Univariate	0.177	0.493	-0.077	-0.131	93.8	90.4
	Multivariate	0.196	0.482	0.096	-0.182	92.2	84.2
	Shared	0.210	0.627	-0.110	-0.327	93.6	86.4
n=200	Marginal	0.177	0.493	-0.077	-0.193	94.5	90.6
	Latent	0.493	0.577	-0.393	0.277	82.6	88.3
	Univariate	0.190	0.483	-0.090	-0.183	94.3	83.2
	Multivariate	0.178	0.490	0.078	-0.190	94.0	92.4
	Shared	0.223	0.608	-0.123	-0.308	90.2	70.5
n=500	Marginal	0.190	0.483	-0.090	-0.183	94.5	83.2
	Latent	0.494	0.551	-0.394	-0.251	55.5	86.9
	Univariate	0.1796	0.489	-0.079	-0.189	92.3	70.9
	Multivariate	0.178	0.490	0.078	-0.190	92.1	80.1
	Shared	0.210	0.613	-0.110	-0.313	92.4	70.9
n=1000	Marginal	0.179	0.489	-0.080	-0.189	92.4	82.2
	Latent	0.482	0.557	-0.382	-0.257	29.6	92
$\rho = 0.8$							
	Univariate	0.173	0.490	-0.073	-0.190	95.4	90.5
	Multivariate	0.180	0.484	0.080	-0.184	93.4	82.2
	Shared	0.606	2.230	-0.506	-1.930	89.3	66.4
n=200	Marginal	0.173	0.490	-0.073	-0.190	95.9	90.3
	Latent	1.057	1.745	-0.957	-1.445	75.8	24.2
	Univariate	0.186	0.482	-0.086	-0.182	94.1	84.3
	Multivariate	0.179	0.489	0.079	-0.189	92.5	92.0
	Shared	0.2222	0.732	-0.533	-0.432	30.4	28.4
n=500	Marginal	0.1018	0.3049	-0.002	-0.005	94.8	90.7
	Latent	0.751	1.204	-0.651	-0.904	43.9	56.1
	Univariate	0.182	0.488	-0.082	-0.188	91.5	70.3
	Multivariate	0.180	0.489	0.080	-0.189	90.3	71.5
	Shared	0.2	0.7086	-0.498	-0.409	12.6	90.8
n=1000	Marginal	0.1013	0.2951	-0.004	0.005	91.8	70.4
	Latent	0.684	1.056	-0.584	-0.756	14.9	90.1

Table 4.6: Estimates, bias and 95% Coverage when analyzing three binary outcomes with different effects. True parameters β_{11}=1.0, β_{12}=0.5, β_{13}=0.3

		Estimates			**Bias**			**95% Coverage**		
Sample size	Model	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{31}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{31}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{31}$
$\rho = \mathbf{0}$										
	Univariate	1.742	0.801	0.516	-0.742	-0.301	-0.216	43.2	81.9	89.2
	Multivariate	2.223	0.344	0.302	-1.223	0.157	-0.002	10.2	32.5	0.1
	Shared R.E.	1.028	0.510	0.311	-0.027	-0.010	-0.011	12.4	66.3	58.2
n=200	Marginal	1.011	0.503	0.307	-0.011	-0.003	-0.007	94.4	95.3	3.5
	Latent	0.871	0.193	-0.05	0.129	0.307	0.351	77.6	91.2	92.3
	Univariate	2.501	1.064	0.506	-1.501	-0.564	0.6	8.7	62.7	81.2
	Multivariate	1.048	0.517	0.267	-0.0478	-0.017	0.0332	30.5	80.5	83.3
	Shared R.E.	1.020	0.502	0.308	-0.020	-0.002	-0.008	90.4	92.2	94.8
n=500	Marginal	1.010	0.497	0.305	-0.010	0.003	-0.005	89.8	92.4	82.9
	Latent	0.855	0.202	-0.046	0.1447	0.2985	0.3464	45.1	87.7	90.4
	Univariate	2.238	1.053	0.510	-1.238	-0.553	-0.2096	0.2	33.5	70.9
	Multivariate	1.081	0.637	0.343	-0.081	-0.137	-0.043	88.4	70.5	92.2
	Shared R.E.	1.009	0.505	0.305	-0.009	-0.005	-0.005	95.3	94.4	93.8
n=1000	Marginal	1.001	0.502	0.303	-0.001	-0.002	-0.003	99.9	99.2	94.1
	Latent	0.849	0.2014	-0.050	0.1514	0.2986	0.3497	14.5	77.1	88.5
$\rho = \mathbf{0.3}$										
	Univariate	4.792	1.336	0.522	-3.792	-0.836	-0.222	46.7	80.3	90.8
	Multivariate	1.414	0.663	0.341	-0.414	-0.163	-0.041	68.4	70.2	82.1
	Shared R.E.	1.229	0.604	0.355	-0.229	-0.104	-0.055	48.5	70.8	72.5
n=200	Marginal	1.009	0.513	0.300	-0.009	-0.013	0.000	77	89.7	90.7
	Latent	1.600	0.744	0.158	-0.134	0.310	0.448	38.5	87.6	92.6
	Univariate	2.531	1.061	0.53	-1.531	-0.561	-0.23	9.8	61.5	80.8
	Multivariate	1.017	0.535	0.373	-0.017	-0.035	-0.073	90.1	84.4	86.5
	Shared R.E.	1.226	0.595	0.358	-0.226	-0.095	-0.057	75.6	83.4	79.8
n=500	Marginal	1.003	0.504	0.302	-0.003	-0.004	-0.002	70.8	73	82
	Latent	1.103	0.194	-0.144	-0.103	0.306	0.444	90.5	81.4	78.3
	Univariate	2.244	1.-399	0.5066	-1.2447	-0.5399	-0.2066	0.3	35.5	72.6
	Multivariate	1.081	0.637	0.343	-0.081	-0.137	-0.043	88.5	70.5	92.2
	Shared R.E.	1.231	0.596	0.357	-0.231	-0.096	-0.057	77.8	83.2	70.5
n=1000	Marginal	1.003	0.502	0.300	-0.003	-0.002	0.000	95.7	94.3	97.5
	Latent	1.095	0.195	-0.144	-0.095	0.305	0.444	83.5	69.8	8.5
$\rho = \mathbf{0.8}$										
	Univariate	5.224	1.351	0.499	-4.224	-0.851	-0.199	46.9	81.9	91.2
	Multivariate	1.803	0.641	0.308	-0.803	-0.141	-0.008	32.3	71.2	93.4
	Shared R.E.	1.506	0.727	0.438	-0.505	-0.227	-0.138	86.4	90.3	91.2
n=200	Marginal	1.004	0.512	0.307	-0.004	-0.012	-0.007	75.5	89.7	91.1
	Latent	2.965	1.455	0.115	-1.965	-0.955	0.185	26.2	70	93.2
	Univariate	2.477	1.077	0.514	-0.757	-0.851	0.520	8.8	59.2	81.7
	Multivariate	1.155	0.642	0.282	-0.155	-0.142	0.018	80.5	70.3	84.3
	Shared R.E.	1.512	0.728	0.431	-0.512	-0.228	-0.131	88.3	73.0	80.2
n=500	Marginal	1.006	0.509	0.302	-0.006	-0.009	-0.002	80.5	71.1	82.7
	Latent	2.916	1.424	0.137	-1.916	-0.924	0.163	0.50	37.2	92.7
	Univariate	2.249	1.043	0.503	-1.249	-0.543	-0.203	0.3	35.5	72.6
	Multivariate	1.132	0.550	0.222	-0.132	-0.050	-0.078	69.5	89.2	40.7
	Shared R.E.	1.505	0.726	0.432	-0.504	-0.226	-0.132	80.0	70.1	86.0
n=1000	Marginal	1.001	0.507	0.302	-0.001	-0.006	-0.002	98.9	90.5	93.8
	Latent	2.878	1.405	0.133	-1.878	-0.905	0.168	0.0	12.0	90.3

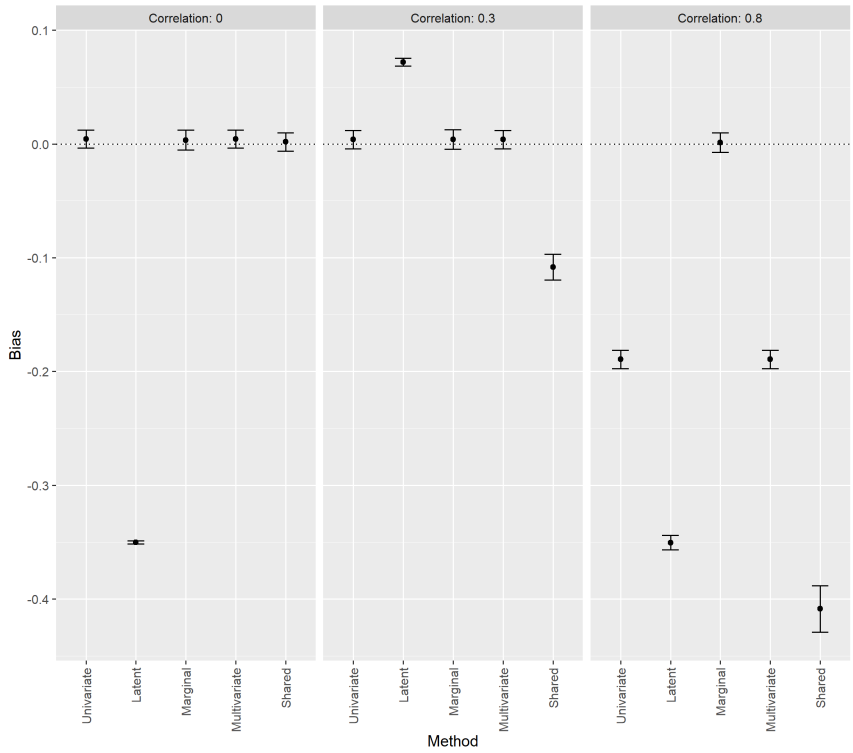


Figure 4.2: Mean treatment effect for two binary outcomes with same effect,$\beta_{\text{binary}} = 0.2$. **Source**: Researcher (2023)

4.5.3 Mixed outcomes

4.5.3.1 Bias

Tables 4.7 and 4.8 provide the estimated effects and bias for two and three mixed outcomes (binary and continuous). The estimates for the univariate, shared random effect and the marginal models were accurate as compared to the estimates for the univariate and latent variable models, see Figure 4.3. The accuracy of the estimates for the marginal and shared random effect models increased when $\rho = 0.8$ and for $n \geq 500$. The accuracy was mostly observed for the estimated effects of the continuous outcomes rather than those for the binary outcomes. The estimates for the latent variable model were less accurate particularly for a small sample size regardless of the correlation strength.

4.5.3.2 95% Coverage

For both scenarios (two or three mixed outcomes), the coverage proportion became closer to the nominal 95% for the marginal mean model for $n = 500$, regardless of the correlation strength.

Table 4.7: Estimates, bias and 95% coverage for two mixed outcomes (one binary and one continuous). True effects $\beta_{11} = 0.5$, $\beta_{21} = 0.3$

		Estimates		**Bias**		**95% Coverage**	
Sample size	Model	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$	$\widehat{\beta}_{11}$	$\widehat{\beta}_{21}$
$\rho = 0$							
	Univariate	0.5005	0.329	-0.0005	-0.1299	96.3	89.8
	Shared	0.5005	0.3299	-5.00E-04	-0.1299	88.2	70.2
n=200	Marginal	0.5005	0.2539	-5.00E-04	-0.0539	96.1	24.1
	Latent	0.410	1.050	-0.0041	0.1234	74.8	8.4
	Univariate	0.502	0.315	-0.0017	-0.1147	95.6	84.5
	Shared	0.5017	0.3147	-0.0017	-0.1147	87.7	14.5
n=500	Marginal	0.5017	0.2458	-0.0017	-0.0458	95.3	1.4
	Latent	0.440	0.040	-0.0024	0.1229	49.9	0.1
	Univariate	0.498	0.323	0.0021	-0.123	94.4	70.3
	Shared	0.4979	0.3232	0.0021	-0.1232	88.2	0.2
n=1000	Marginal	0.4979	0.2495	0.0021	-0.0495	94.4	0.0
	Latent	0.400	0.410	-0.0021	0.121	20.5	0.0
$\rho = 0.3$							
	Univariate	0.4996	0.325	0.0004	-0.125	94.7	90
	Shared	0.4996	0.3246	4.00E-04	-0.1246	87.4	70.4
n=200	Marginal	0.4996	0.2506	4.00E-04	-0.0506	94.3	24.7
	Latent	0.65	1.05	-0.0074	0.1203	65.7	3.7
	Univariate	0.4999	0.3196	0.0001	-0.1196	94.6	84
	Shared	0.4999	0.3196	1.00E-04	-0.1196	89	14.5
n=500	Marginal	0.4999	0.2477	1.00E-04	-0.0477	94.4	0.8
	Latent	0.4	0.22	-0.004	0.1197	38.1	0.0
	Univariate	0.496	0.325	0.0036	-0.125	94.5	69.8
	Shared	0.4964	0.3253	0.0036	-0.1253	94.5	50.5
n=1000	Marginal	0.4964	0.2503	0.0036	-0.0503	94.5	0.0
	Latent	0.48	0.35	-0.0068	0.1172	10.6	0.0
$\rho = 0.8$							
	Univariate	0.499	0.332	0.0007	-0.132	94	90.4
	Shared	0.4993	0.3318	7.00E-04	-0.1318	85.9	71.2
n=200	Marginal	0.4993	0.2544	7.00E-04	-0.0544	93.7	26.3
	Latent	0.69	1.73	-0.0216	0.1046	54.2	0.1
	Univariate	0.499	0.316	0.0008	-0.116	95.9	84.8
	Shared	0.4992	0.3161	8.00E-04	-0.1161	87.2	14.1
n=500	Marginal	0.4992	0.2458	8.00E-04	-0.0458	95.7	1.2
	Latent	0.55	0.88	-0.0181	0.1046	17.7	0.0
	Univariate	0.497	0.326	0.004	-0.126	94.4	69.7
	Shared	0.4965	0.3257	0.0035	-0.1257	85.8	0.5
n=1000	Marginal	0.4965	0.2505	0.0035	-0.0505	94.5	0.0
	Latent	0.52	0.57	-0.0213	0.102	30.1	0.0

Table 4.8: Estimates, bias and 95% Coverage when analysing three mixed outcomes. True parameters β_{11}=0.5, β_{12}=0.2, β_{13}=0.4

		Estimates			**Bias**			**95% Coverage**		
sample size	Model	β_{11}	β_{21}	β_{31}	β_{11}	β_{21}	β_{31}	β_{11}	β_{21}	β_{31}
$\rho = \mathbf{0}$										
	Univariate	0.4922	0.199	0.655	0.0078	0.001	-0.255	94.8	95.7	85
	Shared R.E	0.4922	0.199	0.6553	0.0078	0.001	-0.2553	94.4	96.3	58.2
n=200	Marginal	0.4922	0.4294	-0.1632	0.0078	-0.2294	0.5632	94.4	5.3	3.5
	Latent	0.422	0.245	0.836	0.0996	0.0927	0.3459	87.6	66.4	41.8
	Univariate	0.509	0.1997	0.641	-0.009	0.0003	-0.241	94.2	95.7	72.1
	Shared R.E	0.5091	0.1997	0.6414	-0.0091	3.00E-04	-0.2414	93.9	95.5	30.7
n=500	Marginal	0.5091	0.4429	-0.1602	-0.0091	-0.2429	0.5602	93.9	26.9	0.0
	Latent	0.330	0.298	0.756	0.0825	0.0918	0.3484	0.0	3.7	26.9
	Univariate	0.5007	0.203	0.652	0.0007	-0.003	-0.252	94	96.2	52.7
	Shared R.E	0.5007	0.2031	0.652	-7.00E-04	-0.0031	-0.252	94	96.1	6.3
n=1000	Marginal	0.5007	0.4347	-0.1616	-7.00E-04	-0.2347	0.5616	94	36.1	0.0
	Latent	0.503	0.189	0.407	0.0914	0.0891	0.3462	95	89.3	61.5
$\rho = \mathbf{0.3}$										
	Univariate	0.499	0.204	0.654	0.0015	-0.0041	-0.254	93.8	95.8	86.3
	Shared R.E	0.4985	0.2041	0.6535	0.0015	-0.0041	-0.2535	93.4	95.5	20.1
n=200	Marginal	0.4985	0.4332	-0.1653	0.0015	-0.2332	0.5653	93.4	55.5	2.1
	Latent	0.514	0.214	0.657	0.1316	0.126	0.3847	89.3	65	35
	Univariate	0.506	0.201	0.675	-0.006	0.0007	-0.275	94	96.2	69.6
	Shared R.E	0.5058	0.2007	0.6747	-0.0058	-7.00E-04	-0.2747	94	96.2	20.0
n=500	Marginal	0.5058	0.4413	-0.1614	-0.0058	-0.2413	0.5614	94	36.2	0.0
	Latent	0.535	0.236	0.237	0.128	0.1331	0.3829	72.8	27.2	4.9
	Univariate	0.502	0.198	0.655	-0.002	0.002	-0.255	96	95.4	52.1
	Shared R.E	0.5017	0.198	0.6553	-0.0017	0.002	-0.2553	96	95.4	0.0
n=1000	Marginal	0.5017	0.4375	-0.1619	-0.0017	-0.2375	0.5619	96	55.4	0.0
	Latent	0.514	0.224	0.986	0.1303	0.134	0.3852	93.5	90.3	6.5
$\rho = \mathbf{0.8}$										
	Univariate	0.496	0.2007	0.643	0.004	0.0007	-0.243	93.6	94.8	86.5
	Shared R.E	0.4961	0.2007	0.6429	0.0039	-7.00E-04	-0.2429	93.6	94.2	35
n=200	Marginal	0.4961	0.4319	-0.1648	0.0039	-0.2319	0.5648	93.6	12.1	3
	Latent	0.4498	0.1783	1.529	0.1482	0.1437	0.401	87.9	65	41
	Univariate	0.511	0.206	0.667	-0.011	-0.006	-0.2667	94.1	95.8	72.4
	Shared R.E	0.511	0.206	0.6666	-0.011	-0.006	-0.2666	94.1	95.8	31.5
n=500	Marginal	0.511	0.4437	-0.1614	-0.011	-0.2437	0.5614	94.1	55.8	0.0
	Latent	0.398	0.159	1.865	0.1377	0.1426	0.3996	0.0	31.5	4.6
	Univariate	0.4997	0.1983	0.653	0.0003	0.002	-0.253	94.8	95	53.8
	Shared R.E	0.4997	0.1983	0.653	3.00E-04	0.0017	-0.253	94.7	94.9	63.4
n=1000	Marginal	0.4997	0.4355	-0.1621	3.00E-04	-0.2355	0.5621	94.7	54.9	0.0
	Latent	0.528	0.23	1.827	0.1448	0.1462	0.3981	87.8	65.5	6.0

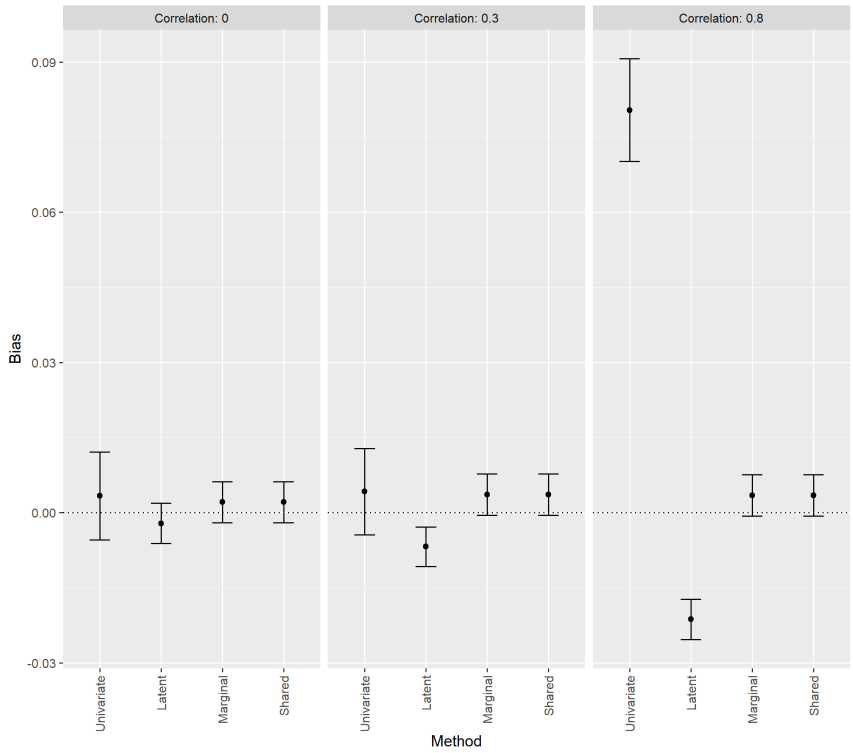


Figure 4.3: Mean treatment effect for two varying mixed outcomes with same effect,$\beta_{\text{continuous}} = 0.5$, $\beta_{\text{binary}} = 0.2$. **Source**: Researcher (2023)

4.5.3.3 Working correlation matrices for GEE

We further assessed the performance of the estimated effects for the marginal mean model using the GEE under different working correlation matrices for correlated binary and continuous outcomes. Tables 4.9 and 4.10 present the accuracy of the estimates under differing correlation strength and working correlation matrix for continuous and binary outcomes, respectively. We observe that there is no difference in the estimates, bias and MSE under the different working correlation matrices regardless of the correlation strength. Therefore, the choice of the working correlation structure chosen in the simulation may not impact the accuracy of the estimates.

Table 4.9: A comparison of estimate, bias and MSE for the marginal model using GEE under different working correlation matrix for continuous outcomes. β_0=-1.1, β_1=0.5

		$\widehat{\beta}_0$			$\widehat{\beta}_1$		
Working correlation matrix	**Correlation**	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
Independence	0	-1.099	-0.001	0.002	0.498	0.001	0.003
	0.3	-1.099	-0.001	0.002	0.497	0.002	0.005
	0.8	-1.099	-0.001	0.002	0.498	0.002	0.005
Exchangeable	0	-1.10	-0.001	0.002	0.499	0.001	0.005
	0.3	-1.099	-0.001	0.002	0.498	0.002	0.005
	0.8	-1.099	-0.001	0.003	0.498	0.002	0.005
Unstructured	0	-1.10	-0.001	0.002	0.499	0.001	0.005
	0.3	-1.099	-0.001	0.003	0.498	0.002	0.005
	0.8	-1.099	-0.001	0.002	0.498	0.002	0.005

Table 4.10: Estimates, bias, MSE for a comparison of the working correlation matrix for estimating effects on binary outcomes using a marginal model. β_0=1.0, β_1=-0.2

		$\widehat{\beta}_0$			$\widehat{\beta}_1$		
Working correlation matrix	**Correlation**	**Estimate**	**Bias**	**MSE**	**Estimate**	**Bias**	**MSE**
Independence	0	1.003	-0.003	0.011	-0.103	0.003	0.02
	0.3	1.004	-0.004	0.011	-0.104	0.004	0.019
	0.8	7.317	-1.823	-6.317	1.623	39.902	2.635
Exchangeable	0	1.003	-0.0005	0.011	-0.103	0.003	0.02
	0.3	1.004	-0.004	0.011	-0.104	0.004	0.019
	0.8	7.317	-1.823	-6.317	1.623	39.902	2.635
Unstructured	0	1.003	-0.003	0.011	-0.103	0.004	0.02
	0.3	1.004	-0.004	0.011	-0.104	0.004	0.019
	0.8	7.317	-1.823	-6.317	1.623	39.902	2.635

4.6 Empirical Study

4.6.1 Data source

We extracted child data from the 2015-16 Malawi Demographic and Health Survey (MDHS). The MDHS is a nationally representative household survey which provides up-to-date information on child health and nutritional status of children under 5 years of age (using weight and height measurements) at national level for both rural and urban areas of the country. The survey followed a stratified two-stage sample design. Parents of infants included in the survey had signed an informed consent. Details of the sampling design for the MDHS can be obtained from the 2015-16 MDHS reports (Government of Malawi & ICF, 2017). The study objective was to estimate the effect of appropriate complementary feeding on height-for-age, weight-for-height, and body-mass-index-for-age (*z-scores*). We controlled for factors that were likely to be associated with appropriate complementary feeding and child growth outcomes such as: place of residence, wealth status, maternal HIV, maternal age, child birth weight, sex of a child and if a child was anemic (Kassa, Meshesha, Haji, & Ebrahim, 2016; Kamenju et al., 2017; Walters et al., 2019).

We assessed the three nutritional outcomes as continuous outcomes, binary outcomes measured as below minus the standard deviation (SD) or above the normal SD, namely; stunting (height-for-age z-score $<$ -2 SD), wasting (weight-for-height < -2 SD) and overweight (body-mass-index-for-age $>$ 2 SD) (WHO, 2006).

To estimate the association between child appropriate complementary feeding and the three child growth outcome measures, namely height-for-age, weight-for-age, and body-mass-

index-for-age simultaneously, we fit the four models. We fit a univariate model to the data for comparison purposes. The parameters for the marginal model were estimated using generalized estimating equation assuming an unstructured working correlation matrix as the correlation between the growth outcomes were different.

4.6.2 Empirical data results

The outcomes WHZ and BAZ are intuitively correlated, but because all three growth measures were indicators of child growth, they could reasonably be assumed to be correlated. Table 4.11 presents estimates for the effect of appropriate complementary feeding on the continuous child growth outcomes for the four models. Appropriate complementary feeding did show to have a positive association on height-for-age, BMI-for-age and, weight-for-height z-scores. However, the associations between appropriate complementary feeding and the growth outcomes were not significant. The estimates of the effects were similar across the methods for height-for-age z-score, however, only the shared R.E (Coef:0.24, 95% CI:(0.01, 0.48)) and latent variable model (Coef:0.30, 95% CI:(0.06, 0.53)) provided significant associations. We also observe that children with low birth weight had decreased height-for-age z-score while female children were more likely to have increased height-for-age z-scores across the four statistical approaches. Additionally, children from rich families had increased weight-for-height z-scores.

Table 4.12 displays estimates for the effect of appropriate complementary feeding on stunting, wasting and overweight. For all methods, We observe that appropriate complementary foods resulted in a decrease in stunting, wasting and overweight, however, the association was not significant. The estimate for the association between appropriate complemen-

Table 4.11: The effect of appropriate complementary feeding on continuous child growth outcomes using univariate, Shared random effect, marginal models

Z-score	Height-for-age	Weight-for-height	BMI for age
	Coef. (95% Conf. Int)	Coef. (95% Conf. Int)	Coef. (95% Conf. Int)
Univariate			
Intercept	-1.44 (-1.95, -0.94)	0.014 (-0.36, 0.39)	0.14 (-0.26, 0.53)
Appropriate (Yes vs No)	0.24 (-0.07, 0.54)	0.08 (-0.16, 0.31)	0.06 (-0.21, 0.33)
Wealth (medium vs poor)	-0.07 (-0.26, 0.13)	0.13 (-0.05, 0.31)	0.16 (-0.02, 0.35)
Wealth (rich vs poor)	0.20 (-0.04, 0.44)	0.22 (0.03, 0.41)	0.19 (-0.01, 0.38)
Mothers HIV (positive vs negative)	-0.14 (-0.49, 0.22)	-0.12 (-0.43, 0.19)	-0.07 (-0.39, 0.25)
Residence (rural vs urban)	-0.24 (-0.56, 0.07)	0.06 (-0.20, 0.32)	0.06 (-0.21, 0.33)
Birth weight ($<$ 2500g vs $\geq$2500g)	-0.19 (-0.45, 0.07)	-0.36 (-0.59, -0.13)	-0.37 (-0.60, -0.13)
Child gender (female vs male)	0.18 (0.02, 0.34)	0.09 (-0.06, 0.23)	0.05 (-0.10, 0.20)
Child anemia (Anemic vs not anemic)	0.01 (-0.18, 0.19)	0.07 (-0.09, 0.23)	0.07 (-0.09, 0.23)
Shared RE			
Intercept	-1.65 (-1.94, -1.35)	-0.05 (-0.34, 0.24)	0.19 (-0.10, 0.49)
Appropriate (Yes vs No)	0.24 (0.01, 0.48)	-0.06 (-0.29, 0.18)	-0.08 (-0.32, 0.15)
Wealth (medium vs poor)	-0.01 (-0.18, 0.16)	0.1 (-0.07, 0.27)	0.14 (-0.04, 0.31)
Wealth (rich vs poor)	0.23 (0.06, 0.39)	0.2 (0.03, 0.36)	0.15 (-0.01, 0.32)
Mothers HIV (positive vs negative)	-0.14 (-0.38, 0.10)	-0.14 (-0.38, 0.11)	-0.12 (-0.36, 0.13)
Residence (rural vs urban)	-0.06 (-0.26, 0.14)	-0.03 (-0.23, 0.16)	0.14 (-0.04, 0.31)
Birth weight ($<$ 2500g vs $\geq$2500g)	-0.27 (-0.50, -0.05)	-0.42 (-0.64, -0.20)	-0.40 (-0.63, -0.18)
Child gender (female vs male)	0.19 (0.06, 0.32)	0.09 (-0.03, 0.22)	0.06 (-0.07, 0.19)
Child anemia (Anemic vs not anemic)	-0.004 (-0.15, 0.14)	0.01 (-0.14, 0.15)	0.002 (-0.14, 0.15)
Marginal			
Intercept	-1.73 (-2.29, -1.18)	0.06 (-0.39, 0.51)	0.30 (-0.17, 0.76)
Appropriate (Yes vs No)	0.24 (-0.04, 0.51)	-0.05 (-0.28, 0.18)	-0.08 (-0.31, 0.16)
Wealth (medium vs poor)	-0.02 (-0.21, 0.17)	0.10 (-0.06, 0.25)	0.12 (-0.04, 0.28)
Wealth (rich vs poor)	0.22 (0.03, 0.41)	0.16 (0.01, 0.31)	0.14 (-0.02, 0.29)
Mothers HIV (positive vs negative)	-0.14 (-0.38, 0.10)	0.13 (-0.37, 0.09)	-0.12 (-0.35, 0.12)
Residence (rural vs urban)	-0.10 (-0.31, 0.13)	-0.10 (-0.28, 0.09)	-0.09 (-0.28, 0.10)
Birth weight ($<$ 2500g vs $\geq$2500g)	-0.27 (-0.52, -0.02)	-0.42 (-0.61, -0.23)	-0.40 (-0.60, -0.21)
Child gender (female vs male)	0.20 (0.06, 0.35)	0.10 (-0.02, 0.21)	0.06 (-0.06, 0.18)
Child anemia (Anemic vs not anemic)	-0.004 (-0.17, 0.16)	0.01 (-0.13, 0.14)	0.002 (-0.14, 0.14)
Latent			
Intercept	-1.19 (-1.57, -0.80)	0.20 (-0.18, 0.58)	0.40 (0.01, 0.78)
Appropriate (Yes vs No)	0.30 (0.06, 0.53)	-0.03 (-0.26, 0.21)	-0.07 (-0.30, 0.17)
Wealth (medium vs poor)	-1.61 (-1.85, -1.37)	0.05 (-0.19, 0.29)	0.28 (0.04, 0.52)
Wealth (rich vs poor)	-1.34 (-1.58, -1.11)	0.15 (-0.09, 0.39)	0.31 (0.08, 0.55)
Mothers HIV (positive vs negative)	-0.16 (-0.40, 0.08)	-0.20 (-0.44, 0.04)	-0.13 (-0.37, 0.11)
Residence (rural vs urban)	-0.23 (-0.41, -0.06)	-0.15 (-0.33, 0.03)	-0.14 (-0.32, 0.04)
Birth weight ($<$2500g vs $\geq$2500g)	-0.28 (-0.50, -0.06)	-0.44 (-0.66, -0.21)	-0.41 (-0.63, -0.19)
Child gender (female vs male)	0.20 (0.07, 0.33)	0.10 (-0.03, 0.23)	0.05 (-0.08, 0.19)
Child anemia (Anemic vs not anemic)	0.03 (-0.11, 0.18)	-0.01 (-0.15, 0.13)	0.01 (-0.13, 0.16)

Conf. Int. - Confidence interval
Coef. - Coefficient

tary feeding and stunting using the univariate model was (OR=0.84, (95% CI:0.51, 1.40)) similar to that of the shared RE (OR=0.82, (95% CI:0.54, 1.25)) and the marginal mean (OR=0.82, (95% CI:0.55, 1.24)) models. For wasting and overweight the effect of appropriate complementary feeding, differed across the models, except for shared RE (OR=0.96, (95% CI:0.37, 2.51)); (OR=0.77, (95% CI:0.34, 1.75)) and the marginal mean model (OR=0.96, (95% CI:0.39, 2.37)); (OR=0.78, (95% CI:0.34, 1.75)) which had similar effects, respectively. We also observe that females were less likely to be stunted across all the models.

For the effect of appropriate complementary foods on the mixed outcomes, appropriate complementary feeding had a positive significant effect on height-for age z-scores using the latent variable model (Coef.=0.29, 95% CI: (0.06, 0.54)). The univariate, marginal mean and shared random effect models showed an increase in height-for-age and weight-for-height z-scores, however, the association was not statistically significant. Children who were given appropriate complementary foods were less likely to be overweight across the methods, however, the association was not statistically significant. The association between appropriate complementary foods and weight-for-height z-scores and overweight were slightly similar across the methods.

4.6.3 Model selection

For the model fit, regression models that used the likelihood to estimate the regression parameters such as the multivariate distribution, shared random effect and the latent variable models we used the Akaike information criteria to select models with the best fit for binary, continuous or mixed outcomes. The ”quasi-likelihood under the independence” model

Table 4.12: The effect of appropriate complementary feeding on binary child growth outcomes using univariate, shared random effect, marginal models

	Stunting	Wasting	Overweight
	Odds Ratio (95% Conf. Int)	Odds Ratio (95% Conf. Int)	Odds Ratio (95% Conf. Int)
Univariate			
Intercept	0.77 (0.40, 1.50)	0.06 (0.01, 0.26)	0.02 (0.01, 0.07)
Appropriate (Yes vs No)	0.84 (0.51, 1.40)	0.73 (0.24, 2.26)	0.54 (0.23, 1.29)
Wealth (medium vs poor)	0.94 (0.68, 1.32)	0.76 (0.36, 1.61)	1.41 (0.67, 2.93)
Wealth (rich vs poor)	0.74 (0.52, 1.04)	0.81 (0.40, 1.67)	1.89 (0.98, 3.66)
Mothers HIV (positive vs negative)	0.92 (0.53, 1.59)	1.67 (0.66, 4.20)	1.26 (0.43, 3.72)
Residence (rural vs urban)	1.41 (0.84, 2.37)	0.69 (0.24, 1.99)	2.65 (1.10, 6.37)
Birth weight ($<$ 2500g vs $\geq$2500g)	1.45 (0.97, 2.16)	1.59 (0.68, 3.72)	0.56 (0.17, 1.84)
Child gender (female vs male)	0.59 (0.46, 0.76)	0.54 (0.29, 1.02)	1.00 (0.58, 1.73)
Child anemia (Anemic vs not anemic)	0.88 (0.65, 1.18)	1.17 (0.65, 2.11)	0.07 (-0.09, 0.23)
Shared RE			
Intercept	0.74 (0.45, 1.21)	0.07 (0.02, 0.22)	0.04 (0.01, 0.11)
Appropriate (Yes vs No)	0.82 (0.54, 1.25)	0.96 (0.37, 2.51)	0.77 (0.34, 1.75)
Wealth (medium vs poor)	0.83 (0.62, 1.11)	0.86 (0.44, 1.68)	1.42 (0.80, 2.51)
Wealth (rich vs poor)	0.68 (0.52, 0.91)	0.66 (0.33, 1.31)	1.7 (1.00, 2.90)
Mothers HIV (positive vs negative)	0.82 (0.53, 1.27)	1.46 (0.63, 3.36)	0.96 (0.43, 2.16)
Residence (rural vs urban)	1.32 (0.92, 1.89)	0.66 (0.31, 1.41)	1.17 (0.63, 2.18)
Birth weight ($<$ 2500g vs $\geq$2500g)	1.65 (1.14, 2.38)	1.30 (0.58, 2.94)	0.63 (0.27, 1.48)
Child gender (female vs male)	0.69 (0.55, 0.86)	0.63 (0.37, 1.06)	0.89 (0.58, 1.36)
Child anemia (Anemic vs not anemic)	0.91 (0.71, 1.18)	1.58 (0.92, 2.72)	0.89 (0.55, 1.44)
Marginal			
Intercept	0.74 (0.46, 1.20)	0.07 (0.02, 0.24)	0.04 (0.02, 0.11)
Appropriate (Yes vs No)	0.82 (0.55, 1.24)	0.96 (0.39, 2.37)	0.78 (0.34, 1.75)
Wealth (medium vs poor)	0.83 (0.63, 1.11)	0.86 (0.44, 1.65)	1.41 (0.80, 2.49)
Wealth (rich vs poor)	0.69 (0.52, 0.910	0.66 (0.33, 1.32)	1.70 (1.00, 2.87)
Mothers HIV (positive vs negative)	0.82 (0.52, 1.45)	1.45 (0.60, 1.50)	0.96 (0.43, 2.52)
Residence (rural vs urban)	1.31 (0.93, 1.87)	0.66 (0.30, 1.45)	1.17 (0.65, 2.13)
Birth weight ($<$ 2500g vs $\geq$2500g)	1.63 (1.01, 1.73)	1.30 (0.59, 1.67)	0.64 (0.27, 1.352)
Child gender (female vs male)	0.70 (0.56, 0.86)	0.63 (0.38, 1.05)	0.89 (0.59, 1.35)
Child anemia (Anemic vs not anemic)	0.92 (0.72, 1.05)	1.58 (0.91, 1.11)	0.89 (0.56, 1.87)
Latent			
Intercept	0.31 (0.16, 0.62)	0.10 (0.02, 0.40)	0.09 (0.03, 0.31)
Appropriate (Yes vs No)	0.77 (0.51, 1.17)	0.85 (0.33, 2.17)	0.83 (0.37, 1.84)
Wealth (medium vs poor)	0.68 (0.46, 0.99)	0.03 (0.01, 0.09)	0.06 (0.03, 0.14)
Wealth (rich vs poor)	0.51 (0.34, 0.76)	0.02 (0.01, 0.08)	0.07 (0.03, 0.16)
Mothers HIV (positive vs negative)	0.82 (0.53, 1.26)	1.99 (0.95, 4.20)	0.92 (0.41, 2.05)
Residence (rural vs urban)	1.69 (1.22, 2.33)	0.92 (0.48, 1.76)	0.88 (0.51, 1.53)
Birth weight ($<$ 2500g vs $\geq$2500g)	1.64 (1.14, 2.360	1.42 (0.66, 3.07)	0.63 (0.27, 1.48)
Child gender (female vs male)	0.69 (0.55, 0.86)	0.60 (0.36, 1.00)	0.91 (0.48, 1.76)
Child anemia (Anemic vs not anemic)	0.89 (0.69, 1.15)	1.47 (0.87, 2.49)	0.91 (0.56, 1.46)

Conf. Int. - Confidence interval

Table 4.13: Estimates for the effect of appropriate complementary foods on HAZ, WHZ and overweight for the different models

	Height-for-age z-score	Weight-for-height z-score	Overweight
	Coef. (95% CI)	Coef. (95% CI)	Odds Ratio (95% CI)
Shared RE			
Intercept	-1.73 (-2.28, -1.18)	0.06 (-0.38, 0.49)	0.04 (0.01, 0.19)
Appropriate (Yes vs No)	0.24 (-0.03, 0.50)	-0.04 (-0.26, 0.17)	0.79 (0.36, 1.74)
Mothers HIV (positive vs negative)	-0.14 (-0.42, 0.14)	-0.13 (-0.35, 0.09)	1.01 (0.46, 2.22)
Residence (rural vs urban)	-0.01 (-0.32, 0.14)	-0.01 (-0.27, 0.08)	1.18 (0.64, 2.18)
Birth weight (< 2500g vs ≥2500g)	-0.27 (-0.53, -0.02)	-0.42 (-0.62, -0.22)	0.63 (0.27, 1.48)
Child gender (female vs male)	0.20 (0.06, 0.35)	0.10 (-0.02, 0.21)	0.89 (0.59, 1.36)
Child anemia (Anemic vs not anemic)	-0.004 (-0.17, 0.16)	0.01 (-0.12, 0.14)	0.88 (0.55, 1.42)
Marginal			
Intercept	-1.69 (-2.25, -1.14)	-0.02 (-0.05, 0.35)	0.03 (0.01, 0.16)
Appropriate (Yes vs No)	0.17 (-0.05, 0.49)	-0.05 (-0.28, 0.18)	0.77 (0.68, 1.35)
Mothers HIV (positive vs negative)	-0.13 (-0.37, 0.10)	-0.13 (-0.38, 0.09)	1.09 (0.69, 1.11)
Residence (rural vs urban)	-0.12 (-0.23, 0.12)	0.04 (-0.18, 0.08)	1.18 (0.97, 1.35)
Birth weight (< 2500g vs ≥2500g)	-0.44 (-0.54, -0.02)	-0.42 (-0.52, -0.01)	0.64 (0.32, 1.24)
Child gender (female vs male)	0.21 (0.04, 0.33)	0.02 (-0.05, 0.20)	0.89 (0.60, 1.34)
Child anemia (Anemic vs not anemic)	-0.001 (-0.15, 0.14)	-0.16 (-0.25, 0.05)	0.89 (0.57, 1.44)
Latent			
Intercept	-1.17 (-1.56, -0.77)	0.23 (-0.17, 0.62)	0.1 (0.03, 0.32)
Appropriate (Yes vs No)	0.29 (0.06, 0.54)	-0.03 (-0.27, 0.21)	0.83 (0.37, 1.84)
Mothers HIV (positive vs negative)	-0.16 (-0.41, 0.10)	-0.19 (-0.44, 0.06)	0.94 (0.42, 2.12)
Residence (rural vs urban)	-0.24 (-0.42, -0.05)	-0.15 (-0.34, 0.03)	0.89 (0.51, 1.54)
Birth weight (< 2500g vs ≥2500g)	-0.28 (-0.51, -0.05)	-0.43 (-0.67, -0.20)	0.63 (0.27, 1.48)
Child gender (female vs male)	0.20 (0.07, 0.34)	0.11 (-0.03, 0.24)	0.91 (0.60, 1.39)
Child anemia (Anemic vs not anemic)	0.03 (-0.12, 0.18)	0.01 (-0.14, 0.16)	0.90 (0.57, 1.46)

Conf. Int. - Confidence interval

criterion (QIC) a semi-parametric approach of model selection was used for the GEE to estimate parameters for the marginal mean model, since the GEE does not specify the likelihood of the joint distribution, (Pan, 2001). Models that had a small AIC were selected as best models fit to the data, similarly for the QIC. Table 4.14 presents results of the AIC and QIC for model selection. For the shared random effects, the latent variable and the marginal mean model, the AIC was small when the multiple outcomes were binary, implying that the models fit well when the outcomes were binary. However, for the multivariate distribution model, the AIC was small when the multiple outcomes were continuous.

Table 4.14: Model selection statistics based on the AIC (QIC)

	Binary	Continuous	Mixed
	AIC (QIC)	AIC (QIC)	AIC (QIC)
Multivariate	3131.89	3121.90	-
Shared R.E	3142	15368	14012
Marginal	(3138.60)	(7676.72)	(5560.0)
Latent Variable	3181.98	15530.53	11410.57

In summary, regardless of the statistical method for multiple outcome used, for our empirical example, the association between appropriate child feeding and child growth outcomes were similar, however, only the shared random effects and the latent variable model showed a significant association between appropriate complementary feeding and height-for-age z-scores for continuous correlated outcomes and when the outcomes were binary and of different scale types.

4.7 Summary

In biomedical studies with several outcomes, some conditions are characterized by these multiple outcomes, and the interest could be centered on estimating treatment effectiveness on them jointly. Therefore, statistical analysis needs to consider the multivariate nature of the data. Multivariate methods have been developed and implemented in statistical software. However, their performance may vary according to the outcome scale type, the number of outcomes, and the pairwise correlation. This chapter has reviewed statistical methods used to assess the association between exposure and multiple outcomes. We mainly evaluated the performance of a marginal, shared random effect and latent variable model in estimating the association of exposure on multiple outcomes. In addition, we have performed a simulation study to investigate the differences in bias and mean square error for different outcome types, sample sizes, and covariate type scenarios. The three methods were then applied to child growth data in Malawi using the Malawi Demographic and Health Survey (MDHS) to assess the effect of appropriate complementary foods on child growth indicators. The simulation results have shown that regardless of the assumed correlations between the outcome measures, the estimates using the shared random effect and marginal models were more accurate for the continuous outcomes. Furthermore, the association estimates were very accurate when estimated using the marginal model under a strong correlation assumption for the binary outcomes and those measured on different scales. In the application, appropriate complementary feeding resulted in better growth of infants and young children. However, since these outcomes may be collected repeatedly over time, to measure the effect of appropriate feeding on child growth, estimating the causal effect would require considering both the between-outcome and the serial correlation. In

the next chapter, we have proposed a method to estimate causal inference for multiple outcomes collected repeatedly over time for observational studies.

CHAPTER 5 : Multiple Causal Effects for Longitudinal Observational data

5.1 Introduction

The causal inference methods discussed so far have concentrated on univariate outcomes, even when we looked at the them in terms of repeated measurements in chapter 3. This Chapter describes proposed methods for the estimation of causal effects for longitudinal studies with multiple outcomes. The proposed methods are centered on extending the marginal structural models with inverse probability treatment for repeated measurements with multiple outcomes.

5.2 Notation

Suppose for subject i, $i = 1, \ldots n$, we observe j outcomes, ($j = 1, \ldots, J_i$), that are repeatedly measured for each subject at time k, $k = 1, \ldots, M_i$. For the purpose of this study and without loss of generality, we set $J_i = J$ and $M_i = M$ $\forall i$ Therefore, each subject will have a JK vector of outcomes that are stacked in vector $\vec{\mathbf{Y}}_i = (Y_{ijk})^T$, where Y_{ijk} denotes the measurement of the j^{th} outcome of subject i at time k. For simplicity we define $\vec{\mathbf{Y}}_i$ as $\mathbf{Y}$. Therefore, the stacked vector can be expanded as $\vec{\mathbf{Y}}_i = (Y_{i11}, \ldots, Y_{i1M}, \ldots, \ldots, Y_{iJ1}, \ldots, Y_{iJM})^T$ = $(\mathbf{Y}_{i1}^T, \mathbf{Y}_{i2}^T, \ldots, \mathbf{Y}_{iJ}^T)^T$, where $\vec{\mathbf{Y}}_i$ is a $JM \times 1$ outcome vector of the i^{th} subject where each of the $\mathbf{Y}_{ij}, (k = 1, \ldots, M)$,

is the j^{th} outcome k-vector measurement of subject i. Suppose at each time point we observe for subject i, treatment T_{ik}, where T_{ik} takes values of 0 and 1. We denote $\bar{T}_{ik} = (T_{i1}, \ldots, T_{iM})$ as the treatment history through time k. We assume for subject i we observe a vector time-varying confounders $\mathbf{X}_{ik}$ at each time point k and $\mathbf{Z} = \mathbf{X}_{i1}$ represents vectors of baseline confounders. We denote $\bar{\mathbf{X}}_{ik} = (X_{i1}, X_{i2}, \ldots, X_{iM})$ as the confounder history through time k for subject i. In addition, we assume a history of outcome-specific confounders defined as $\mathbf{X}_{ijk}$.

We extend the potential outcome framework to the longitudinal multiple outcome setting. Let $\vec{\mathbf{Y}}_i(\bar{T}_k) = (Y_{ijk}(\bar{T}_k))^T$ be the stacked vector of potential outcomes for subject i on the j^{th} outcome measured at time k given that a subject had treatment history $\bar{T}_k$ through time k. Therefore, the average causal effect for subject i is defined as a difference of the average j^{th} potential outcome for the i^{th} subject under different treatment levels at the k^{th} time.

We generalize the definition of the ATE for univariate longitudinal data to a multivariate scenario. The average response vector in the entire population had a subject received $T = 1$ over time is denoted as $E[Y_{ijk}(\bar{T}_k = \bar{1})]$ and the average response vector for the entire population had a subject received $T = 0$ over time is denoted as $E[Y_{ijk}(\bar{T}_k = \bar{1})]$. The ATE for outcome j, repeatedly measured through time k for subject i is given as follows;

$$E[Y_{ijk}(\bar{T}_k = 1)] - E[Y_{ijk}(\bar{T}_k = 0)] \tag{5.1}$$

In vector notation equation 5.1 can be written as;

$$\begin{pmatrix} E(Y_{ij1}(\bar{T}_1 = 1) - Y_{ij1}(\bar{T}_1 = 0)) \\ E(Y_{ij2}(\bar{T}_2 = 1) - Y_{ij2}(\bar{T}_2 = 0)) \\ \vdots \\ E(Y_{ijM}(\bar{T}_M = 1) - Y_{ijM}(\bar{T}_M = 0)) \end{pmatrix} = \begin{pmatrix} E(Y_{ij1}(\bar{T}_1 = 1)) - E(Y_{ij1}(\bar{T}_1 = 0)) \\ E(Y_{ij2}(\bar{T}_2 = 1)) - E(Y_{ij2}(\bar{T}_2 = 0)) \\ \vdots \\ E(Y_{ijM}(\bar{T}_M = 1)) - E(Y_{ijM}(\bar{T}_M = 0)) \end{pmatrix}$$

$$= \mathbf{E}[\mathbf{Y}_{ij}(\bar{T}_k = \bar{1})] - \mathbf{E}[\mathbf{Y}_{ij}(\bar{T}_k = \bar{0})] \tag{5.2}$$

Using vector notation generalizing it for all j outcomes through time k as

$$= \mathbf{E}[\vec{\mathbf{Y}}_i(\bar{T}_k = \bar{1})] - \mathbf{E}[\vec{\mathbf{Y}}_i(\bar{T}_k = \bar{0})] \tag{5.3}$$

For binary outcomes, interest would be to estimate the odds ratio between the treatment groups. If we denote $\boldsymbol{\delta}_{ij}(1) = \frac{\Pr(\mathbf{Y}_{ij}(\bar{T}_k)=\bar{1})}{1-\Pr(\mathbf{Y}_{ij}(\bar{T}_k=\bar{1})=1)}$ as the odds of $\mathbf{Y}_{ij}$ among subjects who had a sequence of treatment and $\boldsymbol{\delta}_{ij}(0) = \frac{Pr(\mathbf{Y}_{ij}(\bar{T}_k=\bar{0})=1)}{1-Pr(\mathbf{Y}_{ij}(\bar{T}_k=\bar{0})=1)}$ as the odds of $\mathbf{Y}_{ij}$ among subjects who had a sequence of no treatment. Then the average treatment effect for the j^{th} binary outcome measured at time k is given as

$$\boldsymbol{\delta}_{OR} = \frac{\boldsymbol{\delta}_{ij}(1)}{\boldsymbol{\delta}_{ij}(0)}$$

5.2.1 Marginal Structural Model for longitudinal data with multiple outcomes

We extend the marginal structural model of equation 3.3 (Hernan et al., 2002) presented in Chapter 3 by incorporating the subscript j for the multiple outcomes measured repeatedly

over time. The effect of exposure history on the vector of marginal means for the j^{th} potential outcome measured at time k for subject i is modeled as follows;

$$E[\mathbf{Y}_i(\bar{T}_k)] = g(h(\bar{T}_k); \boldsymbol{\beta}_j) \tag{5.4}$$

If we consider $h(\bar{T}_k) = \sum_{k=1}^{M} T_k$, then the model in matrix notation for for repeated j continuous outcomes is given as

$$\begin{pmatrix} Y_{ij1} \\ Y_{ij2} \\ \vdots \\ Y_{ijM} \end{pmatrix} = \begin{pmatrix} 1 & \bar{T}_{i1} & 1 \\ 1 & \bar{T}_{i2} & 2 \\ \vdots & \vdots & \\ 1 & \bar{T}_{iM} & M \end{pmatrix} \begin{pmatrix} \beta_{j0} \\ \beta_{j1} \\ \beta_{j2} \end{pmatrix} + \begin{pmatrix} \epsilon_{ij1} \\ \epsilon_{ij2} \\ \vdots \\ \epsilon_{ijM} \end{pmatrix}$$

Which can be generalized as

$$\vec{\mathbf{Y}}_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \text{cum}(\bar{T}_k) + \boldsymbol{\beta}_2 k \tag{5.5}$$

where $\text{cum}(\bar{T}_k) = \sum_{k=1}^{M} T_k$ which is the cumulative exposure and β_{j1} is the causal effect on the j^{th} mean potential outcome $E[\mathbf{Y}_{ij}(\bar{T}_k)]$ for subject i's cumulative exposure.

5.2.2 Assumptions to Identify ATE for multiple outcomes repeatedly measured

We modify the assumptions defined by (Rosenbaum & Rubin, 1983) to account for the longitudinal data with multiple outcomes. The following are sufficient assumptions to identify the causal effect β_{j1} for the j^{th} outcome measured at times $k = 1, \ldots, M$

Assumption 1: Consistency $\mathbf{Y}_{ij}(\bar{T}_k) = \mathbf{Y}_{ij}$

Assumption 2: Positivity; $P(T_k = 1|\mathbf{X}_j, \mathbf{X}_k, \mathbf{Z}) > 0$, each subject has each time probability of being treated and that there is no discontinuation of the treatment. Once an individual is treated, they remain on treatment. We further assume that the treated group follows the same time course as the control group.

Assumption 3: No unmeasured confounder;

$\bar{T}_k \amalg \{\mathbf{Y}_{1j}(T_1), \mathbf{Y}_{2j}(T_2), \ldots, \mathbf{Y}_{Mj}(T_M)\}|\mathbf{X}_j, \mathbf{X}_k, \mathbf{Z}$.

Assumption 4: Stable Unit Treatment Value assumption (SUTVA);

$\{\mathbf{Y}_{1j}(T_1), \mathbf{Y}_{2j}(T_2), \ldots, \mathbf{Y}_{Mj}(T_M)\} \amalg \bar{T}_k$

Assumption 1 implies that a subjects potential outcome vector is equal to the outcome vector for which the actual treatment was received. Assumption 2 addresses overlap and implies that each subject has a non-zero probability of being treated at each time point k. Assumption 3 is the exchangeability, that conditional on the observed covariates, treatment assignment does not depend on the potential outcomes at time k.

The ignorability assumption implies that $P(\mathbf{Y}(T) \leq \mathbf{y}|\mathbf{X}) = P(\mathbf{Y} \leq \mathbf{y}|\mathbf{X})$ (Kennedy, Kangovi, & Mitra, 2019), the conditional distribution of potential outcomes under a certain treatment level $\bar{T}_k = \bar{t}_k$ (given covariates) equals the conditional distribution of observed outcomes (given covariates) among those for whom $T = t$ is observed (Kennedy et al., 2019). The observed outcome is denoted as $\mathbf{Y}_i = \bar{T}_k\mathbf{Y}(1) + (1 - \bar{T}_k)\mathbf{Y}(0)$.

5.2.3 Inverse probability treatment weighting for multivariate longitudinal data : A proposition

We extend the inverse probability treatment weights of equation 3.5 to incorporate the j outcomes. To estimate the longitudinal causal effect with multiple outcomes we incorporate the history of outcome-specific confounders. The pseudo-population generated is one where the average effect of treatment on the multiple outcomes is unconfounded by past exposure, baseline, time-varying, and outcome-specific confounders. We propose modifying the IPTW to account for the history of outcome-specific confounders measured at each time point k known as outcome-specific weights. The unstabilized weight for subject i's j^{th} outcome at time k becomes;

$$W_{ijk} = \prod_{k=1}^{M} \prod_{j=1}^{J} \frac{1}{P(T_k|\bar{T}_{k-1}, \bar{\mathbf{X}}_k, \mathbf{Z}, \bar{\mathbf{X}}_{kj})} \tag{5.6}$$

The estimator in equation 5.5 can be consistently estimated taking into account the modified weights. The stabilized weight becomes

$$SW_{ijk} = \prod_{k=1}^{M} \prod_{j=1}^{J} \frac{P(T_k|\bar{T}_{k-1}), \mathbf{Z}}{P(T_k|\bar{T}_{k-1}, \bar{\mathbf{X}}_k, \mathbf{Z}, \bar{\mathbf{X}}_{kj})} \tag{5.7}$$

Therefore, for multiple outcomes, weighting the observed joint outcomes with the weights equal to the inverse probability of being treated, a pseudo-population is created where the distribution of the confounders is the same. Suppose we define $\pi(\bar{T}_k, \bar{X}) = P(T_k|\bar{T}_{k-1}, \bar{\mathbf{X}}_k, \mathbf{Z}, \bar{\mathbf{X}}_j)$. The average causal effect for the outcome m repeatedly measured across time points k can

be estimated using the IPTW estimator defined as;

$$\widehat{\boldsymbol{\mu}} = \sum_{i=1}^{N}\sum_{j=1}^{J}\frac{T_i Y_{ij}}{\pi(\bar{T}_k, \bar{X}_k)}\left(\sum_{i=1}^{N}\sum_{j=1}^{J}\frac{T_i}{\pi(\bar{T}_k, \bar{X}_k)}\right)^{-1} - \sum_{i=1}^{N}\sum_{j=1}^{J}\frac{(1-T_i)Y_{ij}}{1-\pi(\bar{T}_k, \bar{X}_k)}\left(\sum_{i=1}^{N}\sum_{j=1}^{J}\frac{(1-T_i)}{1-\pi(\bar{T}_k, \bar{X}_k)}\right)^{-1} \tag{5.8}$$

For the above estimator the weights are normalized so that they add up to n_T in each treatment group.

We provide a proof to show that inverse weighting on the observed data estimates the average causal effect. 5.7. Modifying Hirano and Imbens (2001)'s definition of the expectation of the weighted mean under the unconfounded assumption we get;

$$\begin{aligned}
E\left\{\frac{\bar{T}_k \mathbf{Y}_{ij}}{\pi(\bar{T}_k, \bar{\mathbf{X}}), \mathbf{Z}}\right\} &= E\left\{E\left[\frac{\bar{T}_k \mathbf{Y}_{ij}(1)}{\pi(\bar{T}_k, \bar{\mathbf{X}})}\middle| \bar{\mathbf{X}}_k, \bar{\mathbf{X}}_j\right]\right\} \\
&= E\left\{\frac{\mathbf{Y}_{ij}(1)}{\pi(\bar{T}_k, \bar{\mathbf{X}})} E[\bar{T}_k | \bar{\mathbf{X}}_k, \bar{\mathbf{X}}_j, \mathbf{Z}]\right\} \\
&= E\left\{\frac{\mathbf{Y}_{ij}(1)}{\pi(\bar{T}_k, \bar{\mathbf{X}})} P[\bar{T}_k | \bar{\mathbf{X}}_k, \bar{\mathbf{X}}_j, \mathbf{Z}]\right\} \\
&= E\left\{\frac{\mathbf{Y}_{ij}(1)}{\pi(\bar{T}_k, \bar{\mathbf{X}})} P[\bar{T}_k | \bar{\mathbf{X}}_k, \bar{\mathbf{X}}_j, \mathbf{Z}]\right\} \\
&= E\left\{\frac{\mathbf{Y}_{ij}(1)}{\pi(\bar{T}_k, \bar{\mathbf{X}})} * \pi(\bar{T}_k, \bar{\mathbf{X}}_k)\right\} \\
&= E[\mathbf{Y}_{ij}(1)] = E[\mathbf{Y}(1)]
\end{aligned}$$

The same applies for $E\left\{\frac{(1-\bar{T}_k)\mathbf{Y}_{ij}}{(1-\pi(\bar{T}_k, \bar{\mathbf{X}}_k, \mathbf{X}_j, \mathbf{Z}))}\right\} = E(\mathbf{Y}_i(0))$.

5.3 Estimation of the MSM

To estimate the causal parameter in the marginal model 5.5 we adopted the generalized estimating equation approach as it proved to be efficient in estimating the associations of

the exposure on the multiple outcomes in the previous chapter. We proposed a working correlation matrix to account for both the serial and between-outcome correlation within the data. For repeated measures, there is only correlation for repeated measurements within a subject (serial correlation). For repeated measures and multiple outcomes there are both multivariate outcome correlation and between repeated measurements correlation. Using the proposed inverse propensity treatment weights (IPTW), under a correct specification of the exposure model, a consistent estimator can be obtained from model 5.5 by solving the following generalized estimating equations (Robins, 2000)

$$\sum_{i=1}^{n}\sum_{j=1}^{J}\sum_{k=1}^{K}\mathbf{D}_i(\boldsymbol{\beta}).\mathbf{V}^{-1}W_{ijk}(\mathbf{Y}_{ijk}(\bar{t}_k)-\boldsymbol{\mu}_i(\boldsymbol{\beta}))=\mathbf{0} \tag{5.9}$$

$\bar{t}_k$ is the counterfactual exposure history, $\mathbf{D}_i(\boldsymbol{\beta})=\frac{\partial \boldsymbol{\mu}_i}{d\boldsymbol{\beta}}$ is a derivative matrix of the mean vector, where $\boldsymbol{\mu}_i(\boldsymbol{\beta})=(\mathbf{g}(T_1,\mathbf{X};\boldsymbol{\beta}),$

$\mathbf{g}(T_2,\mathbf{X};\boldsymbol{\beta}),\ldots,\mathbf{g}(\bar{T}_{k-1},\mathbf{X};\boldsymbol{\beta}))$. W_{ijk} is the inverse probability treatment weight calculated at each time point for subject i, $\boldsymbol{V}_i$ is the $JM\times JM$ marginal covariance matrix given as

$$\boldsymbol{V}_i=\phi\boldsymbol{\Sigma}_i \tag{5.10}$$

Denoting $\text{Var}(\mathbf{Y}_i)=\phi\boldsymbol{\nu}(\boldsymbol{\mu})$, where $\boldsymbol{\nu}(\boldsymbol{\mu})$ is the variance function and ϕ is the scale parameter. $\mathbf{A}=\sqrt{\boldsymbol{\nu}(\boldsymbol{\mu})}$ is a $JM\times JM$ block diagonal matrix which contains the marginal variance of the outcomes. For binary outcomes, the marginal variance is given as $\boldsymbol{\mu}_i(1-\boldsymbol{\mu}_i)$. $\boldsymbol{\Sigma}_\mathbf{i}=\mathbf{A}^{\frac{1}{2}}\mathbf{R}_i(\boldsymbol{\rho})\mathbf{A}^{\frac{1}{2}}$ and $\mathbf{R}_i(\boldsymbol{\rho})$ is the $JM\times JM$ working covariance matrix for subject i. A total of $\frac{JM(JM+1)}{2}$ unknown parameters $\boldsymbol{\rho}$ are to be estimated from $\mathbf{R}_i(\boldsymbol{\rho})$ which may cause convergence problems when $JM\approxeq N$ (Naik & Rao, 2001; A. Roy &

Khattree, 2005).

To reduce the dimensionality of the parameters to estimate from the working covariance matrix, a parsimonious structure for the working covariance matrix $\mathbf{R}_i(\boldsymbol{\rho})$ through the Kronecker correlation matrix such that (Inan & Yucel, 2017)

$$\mathbf{V}_i = \phi \mathbf{A}^{\frac{1}{2}} (\mathbf{R}_y(\boldsymbol{\tau}) \otimes \mathbf{R}_k(\boldsymbol{\alpha})) \mathbf{A}^{\frac{1}{2}} \tag{5.11}$$

where $\otimes$ is the Kronecker product. We denote $\mathbf{R}_k(\boldsymbol{\alpha})$ as a $M \times M$ working covariance matrix which captures the serial correlation (within-subject) and $\mathbf{R}_y(\boldsymbol{\tau})$ as the $J \times J$ which captures the between-outcome correlation (Naik & Rao, 2001; A. Roy & Khattree, 2005; Inan & Yucel, 2017). We can assume the following structured working correlation to account for the serial correlation; such as exchangeable, $\mathbf{R}(\boldsymbol{\alpha})_{k,k'} = \alpha$, first-order auto-regressive of order one (AR(1)), $\mathbf{R}(\boldsymbol{\alpha})_{k,k'} = \alpha_{k-k'}$, or unstructured, $\mathbf{R}(\boldsymbol{\alpha})_{k,k'} = \alpha_{kk'}$ ($M(M-1)/2$ parameters) for $\mathbf{R}_k(\alpha)$, $\forall k \neq k^{'}$. For the $\mathbf{R}_j(\boldsymbol{\tau})$, an exchangeable: $R(\boldsymbol{\tau})_{j,j'} = \tau$ or unstructured, $R(\boldsymbol{\tau})_{j,j'} = \tau_{j,j'}$, $\forall j \neq j^{'}$, ($J(J-1)/2$ parameters), correlation structures are assumed and can result in a parsimonious model for equation 5.5 (Srivastava, von Rosen, & Von Rosen, 2008; Werner, Jansson, & Stoica, 2008; Brobbey et al., 2021). For example, for correlated binary outcomes repeatedly measured ($j = 2$,$k = 3$) we assumed two Kronecker correlation structures exchangeable between-outcome and Auto-regressive of order one AR(1) between repeated measures correlation. Therefore the working covariance

structure can be written as

$$\mathbf{R}(\boldsymbol{\rho}) = \mathbf{R}_y(\boldsymbol{\tau}) \otimes \mathbf{R}_k(\boldsymbol{\alpha}) = \begin{pmatrix} 1 & \alpha & \alpha^2 & \tau & \tau\alpha & \tau\alpha^2 \\ . & 1 & \alpha & \tau\alpha & \tau & \tau\alpha \\ . & . & 1 & \tau\alpha^2 & \tau\alpha & \tau \\ . & . & . & 1 & \alpha & \alpha \\ . & . & . & . & 1 & \alpha \\ . & . & . & . & . & 1 \end{pmatrix}$$

Therefore we can define the working correlation structure as

$$Corr(Y_{ijk}, Y_{ij'k'}) = \tau\alpha^{|k-k^{'}|},\ \forall\ k \neq k^{'}, j \neq j^{'}$$

$$\mathbf{R}_{jj'kk'} = \begin{cases} 1, & \text{for } k = k^{'}, j = j^{'} \\ \alpha^{|k-k^{'}|}, & \text{for } k \neq k^{'}, j = j^{'} \\ \tau, & \text{for } k = k^{'}, j \neq j^{'} \\ \tau\alpha^{|k-k^{'}|}, & \text{for } k \neq k^{'}, j \neq j^{'} \end{cases}$$

For an assumed exchangeable between-outcome correlation structure and exchangeable within-subject (repeated measures) correlation, the correlation can be defined as

$$Corr(Y_{ijk}, Y_{ij'k'}) = \tau\alpha,\ \forall\ k \neq k^{'}, j \neq j^{'}$$

$$\mathbf{R}_{jj'kk'} = \begin{cases} 1, & \text{for } k = k^{'}, j = j^{'} \\ \alpha, & \text{for } k \neq k^{'}, j = j^{'} \\ \tau, & \text{for } k = k^{'}, j \neq j^{'} \\ \tau\alpha, & \text{for } k \neq k^{'}, j \neq j^{'} \end{cases}$$

Similar definitions can be done for when the between-outcome correlation is unstructured and the correlation between repeated measures is either AR(1) or exchangeable.

Equation 5.9 has been defined for a multivariate potential outcome and we estimate the causal parameters by setting $\bar{T}_k = \{0, 1\}$. Under the consistency and strongly ignorable assumption, $\boldsymbol{\beta}_1$ can be estimated using the observed data and is a consistent estimator by solving the following set of estimating equations

$$\sum_{i=1}^{n}\sum_{j=1}^{J}\mathbf{D}_i(\boldsymbol{\beta}).\mathbf{V}_i^{-1}\widehat{W}_{ijk}(\mathbf{Y}_{ij} - \boldsymbol{\mu}(\boldsymbol{\beta})) = \mathbf{0} \tag{5.12}$$

where $\mathbf{V}^{-1}$ is as defined in equation 5.11, $\widehat{W}_{ik}$ is the inverse probability treatment weights estimated from the data. The GEEs for equation 5.12 can be solved using the fisher scoring algorithm and robust standard errors can be obtained using the sandwich variance estimator. The GEEs of equation 5.12 are solved with a Fisher-scoring algorithm such that

$$\widehat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} + \left(\sum_{i=1}^{n}\tilde{\mathbf{D}}_i(\boldsymbol{\beta})\widehat{\mathbf{V}}_i^{-1}\widehat{W}_{ijk}\tilde{\mathbf{D}}_i(\boldsymbol{\beta})\right)^{-1}\left(\sum_{i=1}^{n}\tilde{\mathbf{D}}_i(\boldsymbol{\beta})\widehat{\mathbf{V}}_i^{-1}\widehat{W}_{ijk}(\mathbf{Y}_i - \tilde{\boldsymbol{\mu}}_i)\right) \tag{5.13}$$

and a sandwich formula to estimate the asymptotic covariance matrix of the GEE estimators are given as follows;

$$\text{Var}(\widehat{\boldsymbol{\beta}}) = \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}\widehat{\mathbf{A}}^{-1} \tag{5.14}$$

where $\widehat{\mathbf{A}} = \left(\sum_{i=1}^{n}\widehat{\mathbf{D}}_i^T\widehat{\mathbf{V}}_i^{-1}\widehat{W}_{ijk}\widehat{\mathbf{D}}_i\right)^{-1}$ and $\widehat{\mathbf{B}} = \left(\sum_{i=1}^{n}\widehat{\mathbf{D}}_i^T\widehat{\mathbf{V}}_i^{-1}\widehat{W}_{ijk}\widehat{\text{Cov}}(\mathbf{Y}_i)\widehat{\mathbf{V}}_i^{-1}\widehat{W}_{ik}\widehat{\mathbf{D}}_i\right)^{-1}$
and $\text{Cov}(\mathbf{Y}_i) = (\mathbf{Y}_i - \widehat{\widehat{\boldsymbol{\mu}}})(\mathbf{Y}_i - \widehat{\widehat{\boldsymbol{\mu}}})^T$

5.4 Application: Malawi Longitudinal Study of Families and Health

In this example, the interest is to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners ($j = 1, 2$) reported at each year of study 2004, 2006, 2008,

2010, and 2012, 2013, 2017, 2018 and 2019 ($k = 0, 1, 2, 3, \ldots, 8$).The marginal structural model for multiple outcomes was fitted using a logit link function from the MLSFH for the two binary outcomes as a function of HIV status awareness and time under the strongly ignorable assumption as follows:

$$\text{logit}(\Pr(Y_{ijk} = 1|\text{HIV awareness}, k)) = \beta_{1j}\text{HIV awareness}_i + \beta_{2j}k \tag{5.15}$$

where Y_{ijk} is the the j^{th} outcome, $j = 1$ for condom use and $j = 2$ for two or more partners, of the i^{th} individual at time k. We assumed two Kronecker correlation structures and denoted the the serial correlation structure as $R_k(\boldsymbol{\alpha})$ and the between-outcome correlation structure as $R_j(\boldsymbol{\tau})$. The causal parameters from the marginal model in equation 5.15 were estimated under the following assumed working correlation structures combination; exchangeable between-outcome and exchangeable serial correlation structures; exchangeable between-outcome and AR(1) serial correlation structures; unstructured between-outcome and exchangeable serial correlation structures; unstructured between-outcome and AR(1) serial correlation structures, see Table 5.3. Only one correlation parameter was estimated for the unstructured correlation structure ($J(J-1)/2$). We fitted the multivariate model taking into account the inverse propensity treatment weights to control for confounding. We estimated the stabilized weights by first fitting a binary logistic regression model for exposure HIV positivity awareness as a function of all baseline, time-varying, and outcome-specific confounders for each outcome using the `IPW` package in R. A partner's HIV status knowledge was considered as the outcome-specific confounder for condom use, and self-efficacy for multiple sexual partners. The final weight was obtained as a product of outcome-specific weights.

We assessed the statistical approach in terms of the association with and without stabilized weights for comparison purposes, where the dependence between outcomes is modelled by specifying a working correlation structure. The causal parameter for the proposed methods was estimated using the joint generalized estimating equation solver package `JGEE` in R where both working correlation structures are specified (Inan & Yucel, 2017). We modified the `JGEE` source code to incorporate the stabilized weights in the estimation as initially, this was undefined.

5.4.1 Results

Table 5.1: Estimates for a bivariate logistic regression model without IPTW

	Odds ratio	95% Conf. Int	Std.Error
Outcome: Condom use			
Intercept	0.19	(0.18, 0.21)	0.03
HIV aware (HIV negativity vs HIV positivity)	2.53**	(2.05, 3.11)	0.11
Gender (women vs men)	1.55**	(1.37, 1.76)	0.07
Age (in years)	0.95**	(0.94, 0.96)	0.003
Wealth (poor vs medium/rich)	1.03	(0.91, 1.17)	0.07
Education level (years)	1.32**	(1.09, 1.59)	0.10
Religion (Muslim or other vs Christian)	1.04	(0.91, 1.19)	0.07
Region (North vs Central)	0.67	(0.44, 1.02)	0.22
Region (North vs South)	1.21	(0.80, 1.85)	0.23
Wave	0.87	(0.85, 0.89)	0.01
Outcome: Two or more partners			
Intercept	0.25	(0.23, 0.27)	0.04
HIV aware (HIV negativity vs positivity)	1.42**	(1.13, 1.77)	0.12
Gender (women vs men)	4.74**	(4.17, 5.39)	0.07
Age (in years)	0.97**	(0.96, 0.98)	0.002
Wealth (poor vs medium/rich)	0.79**	(0.71, 0.88)	0.05
Education level (years)	0.72**	(0.62, 0.83)	0.08
Religion (Muslim or other vs Christian)	0.85**	(0.76, 0.96)	0.06
Region (North vs Central)	1.31**	(1.09, 1.56)	0.09
Region (North vs South)	1.05	(0.87, 1.26)	0.10
Wave	0.68	(0.66, 0.70)	0.02

** p-value <0.001

Conf. Int. - Confidence interval

SE - Standard error

5.4.2 *Generation of IPTWs for multiple outcomes*

The stabilized weights for each outcome j were generated as follows;

$$SW_{ijt} = \prod_{t=0}^{8} \prod_{j=1}^{2} \frac{P(A_t = 1|\bar{A}_{t-1}, \mathbf{X}_0)}{P(A_t = 1|\bar{A}_{t-1}, \bar{\mathbf{X}}_t, \mathbf{X}_0, \bar{\mathbf{X}}_j)}$$

To account for drop-out (refused to participate, dead, moved out or not eligible) we generated the dropout weights

$$CW_{it} = \prod_{t=0}^{8} \frac{P(\text{Censored}_t = 0|\bar{\text{Censored}}_{t-1} = 0, \bar{A}_t, \mathbf{X}_0)}{P(\text{Censored}_t = 0|\bar{\text{Censored}}_{t-1} = 0, \bar{A}_t, \bar{\mathbf{X}}_t, \mathbf{X}_0, \bar{\mathbf{X}}_j)}$$

The final weight $SW_i = SW_{ijt} \times CW_{it}$

Table 5.2: Distribution of inverse probability treatment weights

	Mean (SD)	**Median**	**Interquartile range**	**Min**	**Max**
Stabilized	1.01 (0.13)	1.01	(0.99, 1.01)	0.1	2.60
LTFU weights	0.99 (0.07)	1.00	(0.99, 1.01)	0.33	1.85
Final weights	1.00 (0.14)	1.00	(0.99, 1.01)	0.1	2.53

5.4.3 *Proposed MSM*

Table 5.3 presents estimates for the proposed methods (bivariate MSM-IPTWs). We observe from Table 5.3 a positive effect of HIV positivity awareness on condom use. Individuals aware of their HIV positivity were more likely to use condoms than those aware of their HIV negativity. There was no significant causal association between awareness of HIV positivity and multiple sexual partners. Furthermore, the estimates obtained using the stabilized weights differed significantly from those obtained without the weights.

The difference was mainly observed when the assumed working correlation structure was unstructured and exchangeable for the between-outcome correlation and first-order autoregressive (AR(1)) for the within-subject correlation. The estimates for the causal effects had smaller standard errors when the IPTW were considered, even though the estimates were slightly similar to those when control for confounding using IPTW is ignored. Furthermore, a comparison of the causal estimates for the univariate causal methods to the proposed methods has shown that the estimated standard errors and awareness of HIV positivity effects were smaller for the proposed methods resulting in a more significant causal association than those obtained using the univariate methods.

5.4.4 Sensitivity analysis

The causal associations for the proposed methods after considering the stabilized weights presented in Table 5.3 have been obtained under the assumption of no unmeasured confounders. (VanderWeele et al., 2020) suggested the use of the E-value approach when dealing with multiple outcomes that are repeatedly measured. Therefore, we assessed the robustness of the estimated causal associations from unmeasured confounding using the E-value statistics for an estimated odds ratio. Table 5.4 presented the E-value statistics for the point estimates of the causal associations when the within-subject correlation was assumed to be AR(1), and the between-outcome correlation was assumed to be exchangeable as shown in Table 5.3. We observed that the estimated causal effcet for condom use was more robust to unmeasured confounding than the estimates for multiple sexual partners.

Table 5.3: HIV status awareness effects on condom use and multiple sexual partners using the joint MSM with IPTW

$R_y(\tau)$	Exchangeable				Unstructured			
$R_t(\alpha)$	Exchangeable		AR(1)		Exchangeable		AR(1)	
	Odds ratio (SE)	95% Conf. Int	Odds ratio (SE)	95% Conf. Int	Odds ratio (SE)	95% Conf. Int	Odds ratio (SE)	95% Conf. Int
Condom use								
HIV negativity	1.00		1.00		1.00		1.00	
HIV positivity	1.95 (0.10)	(1.63, 2.33)	1.90 (0.09)	(1.59, 2.27)	1.95 (0.10)	(1.63, 2.33)	1.89 (0.09)	(1.58, 2.33)
Multiple partners								
HIV negativity	1.00		1.00		1.00		1.00	
HIV positivity	0.94 (0.10)	(0.77, 1.16)	0.95 (0.10)	(0.78, 1.16)	0.94 (0.09)	(0.77, 1.15)	0.95 (0.10)	(0.78, 1.16)

** p-value <0.001
Conf. Int. - Confidence interval

Table 5.4: E-value to assess robustness to unmeasured confounding for the causal effect of HIV status awareness on condom use and multiple partners

Outcome	**E-value for point estimate**	**E-value for LL of CI**
Condom use		
HIV positivity	2.16	1.85
Two or more partners		
HIV positivity	1.48	1.00

LL - Lower limit

CI - Confidence Interval

E-value parameters calculated as a function of the adjusted odds ratio (OR) for the effect of HIV positivity awareness on the outcomes

Furthermore, a sensitivity analysis was also conducted to assess the influence of dropping the confounders in changing the estimated effect. From Figure 5.1, we observe that dropping age in years, years of education, and all the confounders at once would result in a change in the estimated effect of awareness of HIV positivity on condom use. If all covariates were dropped it would change the effect drastically. Hence, it was important to include these variables in the IPTWs.

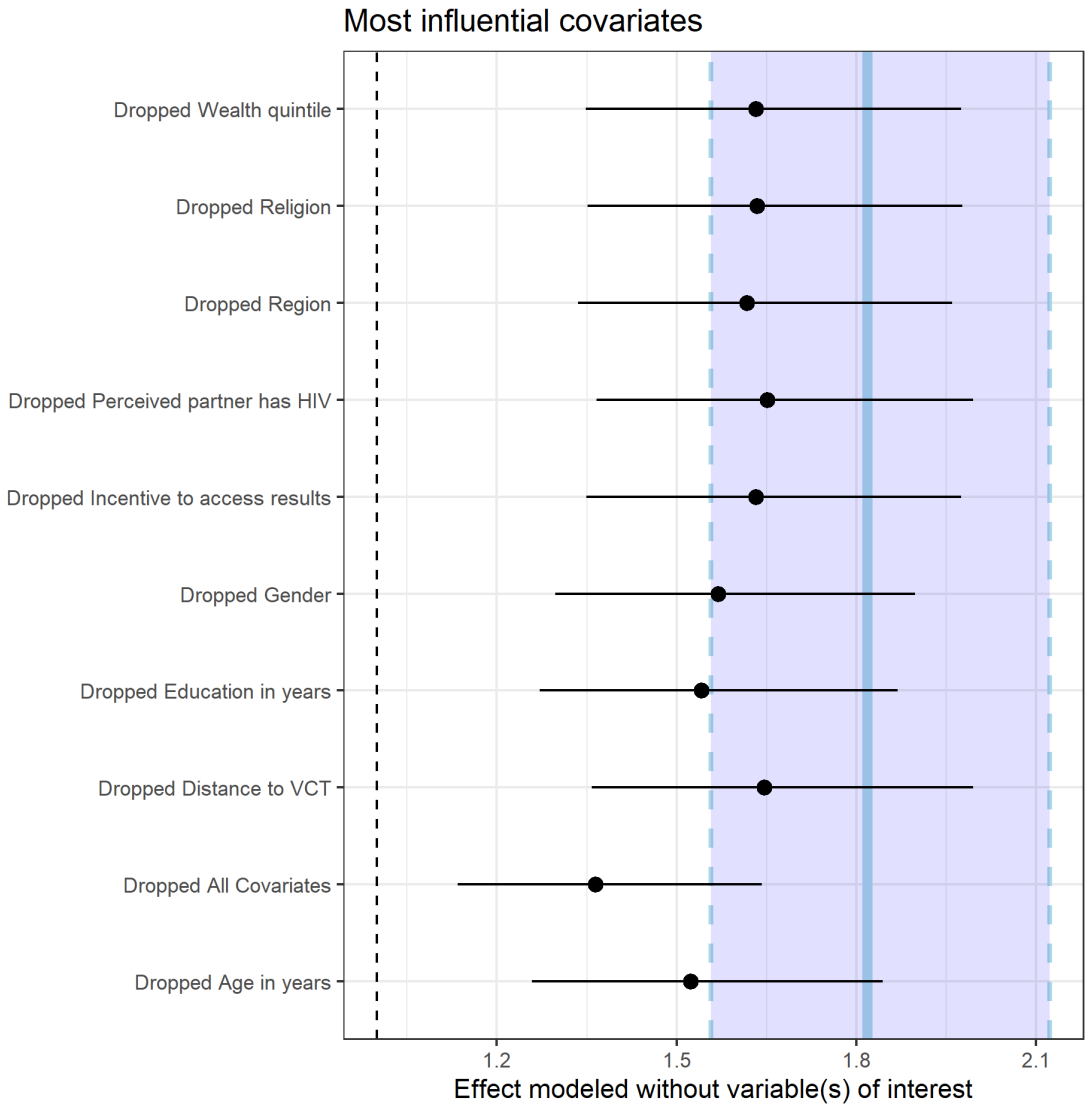


Figure 5.1: Confounder influence on the estimated effect of HIV positivity awareness on condom use. **Source:** Researcher (2023)

5.5 Summary

This chapter has proposed a statistical approach to estimate treatment effect for longitudinal observational data with multiple outcomes. The proposed method takes into account the within and between-outcome correlations. A modified inverse probability treatment weights have been used to control for time-varying confounders, the exposure history,

and outcome-specific confounders. We have illustrated the application of the proposed method on the Malawi Longitudinal Study of Families and Health to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners. We find that HIV positivity awareness is associated with condom use. However, using the proposed methods the estimates of the standard errors were smaller compared to the standard univariate longitudinal models. The proposed method can be implemented using available software with a slight modification to incorporate the weights.

CHAPTER 6 : Discussion and Conclusion

This Ph.D. research has studied causal inference methods that could be used to estimate treatment effects in longitudinal observational studies with multiple outcomes. The studied methods go beyond the recent development in univariate causal inference for longitudinal studies. Our proposed methods have been built around the Marginal Structural Models with inverse probability treatment weights (MSM-IPTWs). The IPTWs have been re-defined as a product of the inverse weights at each time point. The IPTWs have been generated by modelling treatment as a function of prior treatments, baseline, time-varying and outcome-specific confounders. The effect estimates for the MSM-IPTWs have been estimated using generalised estimating equations and accounted for the dependency within and between the outcomes. The method's utility has been demonstrated through an application to estimate the effect of HIV positivity awareness on condom use and multiple sexual partners using the Malawi Longitudinal Study of Families and Health (MLSFH).

The results have shown that awareness of HIV positivity was associated with condom use and multiple sexual partners. Individuals aware of their HIV positivity are more likely to use a condom and have reduced multiple sexual partners. Both estimates of the effects and their associated standard were smaller under the proposed methods than the standard bivariate methods. However, the actual substantive effects were not affected as presented in a recent study for univariate longitudinal outcomes

(Twabi, Manda, Small, & Kohler, 2022). These results of the effect of HIV positivity awareness on sexual behaviour have been found in public health literature (Jewkes et al., 2006; Fedor et al., 2015; Anglewicz, 2012; Anglewicz & Clark, 2013; Adedimeji et al., 2015; George et al., 2019). Thus, this could indicate that our proposed methods correctly adjusted for the confounders and the dependency structure within and between the outcomes. Some studies have shown that using standard covariate adjustment methods may overestimate treatment effects when dealing with longitudinal data as covariate adjustment can block the indirect effect of time-varying exposures on the outcomes (Robins, 1997; Robins, Greenland, & Hu, 1999; Imai & Ratkovic, 2015; VanderWeele et al., 2020).

IPTWs are known to appropriately adjust for confounding and consistently estimate the treatment exposure effects of MSMs under the ignorability assumption (exchangeability, no unmeasured confounders and positivity) (Robins, 1997; Hernan et al., 2000; Robins, 2000; Hernan et al., 2002; Cole & Hernán, 2008; Robins & Hernan, 2009). For our proposed methods, the IPTWs have been used to control for time-varying treatments and confounders (Robins, 1997; VanderWeele, 2009; Naimi et al., 2017). Estimating treatment exposure effects for multiple outcomes using a marginal structural model with inverse probability treatment weights has been done in (Moodie et al., 2014) and (Agogo et al., 2019). Moodie et al. (2014) used a semi-parametric survival method to obtain cause-specific hazard ratios and did not specifically capture the complex dependency structure. Agogo et al. (2019) used a shared random effect model with IPTWs to jointly model binary longitudinal outcomes. However, they did not account for time-varying and outcome-specific confounders. In both studies, simulations were performed and showed the efficiency of the IPTWs in estimating treatment exposure effects for multiple outcomes. Thus, our proposed methods

are suitable for estimating exposure effects as we combine methods that account for both confounding and the dependency structure within and between the outcomes. Since, we used the MSM-IPTWs under the assumption that there are no unmeasured (which is not testable), it is important to conduct sensitivity analyses tests to explore the impact of unobserved confounding or other sources of bias that can distort the estimated treatment effects. Through sensitivity analyses one can better understand the robustness of the causal conclusions and provide more reliable and informative results (Brumback, He, Prasad, Freeman, & Rheingans, 2014; VanderWeele & Ding, 2017; Rosenbaum, 2002). Apart from the IPTWs used to control for confounding, the instrumental variable (IV) approach is one option for controlling for measured and unmeasured confounders. It involves by finding a third variable that is directly associated with the exposure and indirectly associated with the outcome through the exposure but not associated with any confounders. By satisfying certain assumptions, the IV approach can estimate the causal effect of the exposure on the outcome while controlling for confounding (Sainani, 2018).

For longitudinal data with multiple outcomes, it is difficult to specify the full joint likelihood to estimate treatment effects. Thus, generalised estimating equations are commonly used (Liang & Zeger, 1986) as they are also computationally efficient when the multiple outcomes are non-normal (Liang et al., 1992; Zeng & Cook, 2007). For our proposed methods, the dependency between and within the outcomes has been accounted for by specifying a working correlation matrix using a Kronecker product to reduce the number of covariance parameters to estimate. However, studies found that the Kronecker product may be restrictive and sensitive to misspecification. It may also be computationally intensive to estimate the parameters for an unstructured correlation structure (Lu, Xue, &

Hu, 2020; Verbeke et al., 2014). Cho (2016) used a quadratic inference function (QIF) to improve the efficiency of estimating parameters for multivariate marginal models. One of the advantages of the QIF is that it can be used for multiple outcomes of different scale types. Furthermore, it readily accounts for the complex correlation structure without estimating the correlation parameters and specifying the likelihood functions. Even though the Kronecker product may have some limitations, studies have shown that multivariate regression methods that account for the complex dependency structure for multivariate longitudinal studies provide more accurate and precise estimates than methods that treat each outcome separately (Verbeke et al., 2014; Cho, 2016; Inan & Yucel, 2017; Fieuws & Verbeke, 2006).

Most marginal structural models assume that the treatment exposure is fixed and does not change over time (Robins, 1997; VanderWeele, 2009; VanderWeele et al., 2016). For the application of our proposed methods on the HIV behavioural study, an individual may not be aware of their positive HIV status. Once aware of their HIV positivity, they remain aware of a positive HIV status. Our proposed methods specified the treatment exposure history as a total awareness of HIV positivity over time. The estimate quantified the cumulative exposure effect (linear combination of awareness of HIV positivity at different time points) on the outcomes. Most studies that have used the MSM have considered the most recent treatment exposure status on the outcome at the end of a study period (Tsai et al., 2010; Yang et al., 2015; Danaei, Rodríguez, Cantero, Logan, & Hernán, 2013). Most of these studies have used the IPTWs to control for time-varying confounders to avoid blocking the indirect effects of the exposure. The challenge with using the recent exposure while using stabilised IPTWs to control for past confounders may bias the exposure effects because

the status of the previous exposure is not balanced between the recent treatment exposure groups (Lipkovich, Mallinckrodt, & Faries, 2012). The exposure plan can be specified as an estimate of the average effect on the outcomes in the follow-up period. It can further be specified as never treated, always treated, treated every other time point, or treated after four-time points (Platt, Brookhart, Cole, Westreich, & Schisterman, 2013; Hernan et al., 2000; Robins, 2000).

Marginal or conditional models are commonly used to model multivariate longitudinal data (Brobbey et al., 2021). In these models, confounder control is mainly done through covariate adjustment. However, covariate adjustment could be limited when the outcome model is misspecified or when the covariates adjusted are non-linear (Kahlert et al., 2017), hence, the presence of residual confounding. For our proposed methods, we used a marginal model for the potential outcomes with IPTWs to estimate the average treatment effects of HIV positivity awareness and control for confounding. Standard statistical balancing techniques include propensity score techniques (matching, weighting, subclassification) (Austin & Mamdani, 2006; Austin, 2008; Stuart, 2010; Austin, 2014; Austin & Stuart, 2017) could also be used to control for confounding at each time point. Other balancing approaches that could have been used include calibrated inverse probability weights and instrumental variable approaches that control for both measured and unmeasured confounders (Tan, 2010b; Kalia, Saarela, Escobar, Moineddin, & Greiver, 2023; Michael, Cui, Lorch, & Tchetgen, 2020; Cui, Michael, Tanser, & Tchetgen Tchetgen, 2023). For the MSM-IPTWs in our proposed methods, the selection of important confounder variables to include in the treatment model could have been done using machine learning techniques, such as boosted regression (McCaffrey et al., 2013), random forests and neural networks

(McConnell & Lindner, 2019; Karim, Platt, & Group, 2017). The IPTWs we have used for our proposed methods are prone to misspecification if the treatment model is not correctly specified (Cole & Hernán, 2008). In recent studies, doubly robust weighted techniques are used to avoid misspecification of the treatment model (Tan, 2010a; Funk et al., 2011). The approach generates weights that specify both the outcome and treatment model, and either should be correctly specified (Cao et al., 2009; Tu, Koh, & Jiao, 2013).

Our proposed methods have used a marginal modelling approach, where a marginal mean model has been specified for each outcome. The causal effects have been estimated using weighted estimating equations, and the dependency between and within the outcomes has been accounted for through a working correlation matrix (J. Roy, Lin, & Ryan, 2003). However, conditional models such as the generalised linear mixed model, where the correlation between the outcomes is captured using unobserved random effects, could be used in estimating effects as in Agogo et al. (2019). One of the challenges with using the generalised linear mixed-effects models is that the full likelihood should be specified (Laird & Ware, 1982). As the number of multiple outcomes increases, the number of random effects may increase if the mixed-effects model does not assume a shared unobserved factor between the outcomes. Thus, the random-effects methods are more likely to face computational challenges (Cho, 2016; Inan & Yucel, 2017). One could jointly model the outcomes using a pairwise joint modelling approach where the joint models for all outcome pairs are fit and combined (Fieuws & Verbeke, 2006; Bai, Zhong, Gao, & Xu, 2020).

6.1 Strengths and Limitations of the study

Our proposed methods of integrating marginal structural models and inverse probability treatment weights (MSM-IPTWs) have been used to estimate HIV positivity awareness on condom use and multiple sexual partners. The proposed methods could be implemented in statistical software with a bit of adjustment. Unlike previous studies (Hernan et al., 2000, 2002; Tsai et al., 2010; Imai & Ratkovic, 2015; Bodnar, Davidian, Siega-Riz, & Tsiatis, 2004; VanderWeele, 2017; VanderWeele et al., 2016), our proposed method has integrated three statistical approaches that involve time-varying effects, time-varying confounders, the possible dependency between and within multiple outcomes in one single model. The IPTWs have been used to control for confounders in estimating the effects under the assumption of no unmeasured confounders. However, not all confounders were measured, which could be a limitation of the results. Furthermore, some of the important data were missing, and we did not account for the missingness in the analysis. This could be another limitation to the results of our proposed methods. Regarding methodology limitations, using the IPTWs was subjective as it only controls for measured confounders. The treatment model used to generate the weights and propensity scores can be prone to misspecification if some variables, including interactions, are omitted. Our proposed method employed a Kronecker product to account for the between and within outcome correlations. However, other approaches that account for the correlation using unobserved random effects through a generalised multivariate linear mixed model could have been used.

6.2 Future research directions

Future studies of similar methods could look at different balancing options, for example, propensity score matching, generalised propensity scores, doubly robust weighting methods, and instrumental variable approaches. In terms of estimating the treatment exposure effects, the G-estimation, G-computation and Bayesian approaches methods(Vansteelandt, Joffe, et al., 2014; Dukes & Vansteelandt, 2018; Robins, 2000; Picciotto & Neophytou, 2016; Naimi et al., 2017) could be used. Further work could look at different correlation structures, for example, where the serial correlation can depend on the distribution of the data points and the cross-correlation between outcomes at different time points.

APPENDIX

A Estimates of HIV positivity awareness effects using Unstabilized weights

Table 1: Estimates of the effect of awareness of HIV positivity on risky sexual behavior using unstabilized weights

(a) Effect on condom use

Awareness of HIV status	**Odds ratio**	**95% CI**	**p-value**
HIV negativity	1.00		
HIV positivity	2.87	(1.67, 4.95)	<0.001

(b) Effect on multiple sexual partners

Awareness of HIV status	**Odds ratio**	**95% CI**	**p-value**
HIV negativity	1.00		
HIV positivity	0.94	(0.48, 1.84)	0.85

95% Confidence Interval

Table 2: Estimates, bias and 95% Coverage for the exposure effect on three continuous outcomes with same covariates. β_0=-1.1, β_1=0.5

Sample size	Model	$\widehat{\beta}_0$	$\widehat{\beta}_1$	Bias ($\widehat{\beta}_0$)	Bias ($\widehat{\beta}_1$)	95% Cov ($\widehat{\beta}_0$)	95% Cov ($\widehat{\beta}_1$)
ρ**=0**							
	Univariate	-1.097	0.498	-0.0035	0.0018	94.8	95.7
	Multivariate	-1.102	0.502	0.0016	-0.0016	93.2	95.5
	Shared	-1.098	0.499	-0.002	0.002	94.4	95.3
n=200	Marginal	-1.098	0.499	-0.002	0.002	94.4	95.3
	Latent	-1.098	0.316	-0.002	0.184	53.8	50.4
	Univariate	-1.105	0.509	0.0047	-0.0094	94.2	95.7
	Multivariate	-1.105	0.509	0.0047	-0.0094	94.1	95.7
	Shared	-1.000	0.500	-0.0003	-0.0003	93.9	95.5
n=500	Marginal	-1.100	0.500	-0.0003	-0.0003	93.9	95.5
	Latent	-1.100	0.314	0.000	0.186	15.7	12.5
	Univariate	-1.1005	0.5005	0.0005	-0.0005	96.2	95.7
	Multivariate	-1.1005	0.5005	0.0005	-0.0005	96.0	95.7
	Shared	-1.101	0.502	0.0007	-0.002	94	96.1
n=1000	Marginal	-1.101	0.502	0.0007	-0.002	94	96.1
	Latent	1.101	0.315	0.001	1.0	0.7	0.9
ρ**=0.3**							
	Univariate	-1.097	0.496	-0.0034	0.0041	93.8	95.8
	Multivariate	-1.097	0.496	-0.0034	0.0041	93.5	94.4
	Shared	-1.099	0.499	-0.001	0.001	93.4	95.5
n=200	Marginal	-1.099	0.499	-0.001	0.001	93.4	95.5
	Latent	-1.099	0.258	-0.001	0.242	94.5	49
	Univariate	-1.1054	0.511	0.0054	-0.011	94	96.2
	Multivariate	-1.104	0.509	0.0035	-0.0086	93.6	91.4
	Shared	-1.101	0.504	0.001	-0.004	96.2	94.3
n=500	Marginal	-1.101	0.504	0.001	-0.004	96.2	94.3
	Latent	-1.101	0.259	0.001	0.241	90.2	11.0
	Univariate	-1.0995	0.5014	-0.0005	-0.0014	96	95.4
	Multivariate	-1.0995	0.501	-0.0005	-0.001	95.2	90.4
	Shared	-1.099	0.499	-9.00E-04	8.00E-04	96	95.4
n=1000	Marginal	-1.099	0.499	-9.00E-04	8.00E-04	96	95.4
	Latent	-1.099	0.257	0.001	0.244	95.8	0.1
$\rho = 0.8$							
	Univariate	-1.100	0.500	0.0002	-0.0002	93.6	94.8
	Multivariate	-1.100	0.500	0.0002	-0.0002	97.1	96.6
	Shared	-1.099	0.499	-0.002	0.002	93.6	94.4
n=200	Marginal	-1.099	0.499	-0.002	0.002	93.6	94.4
	Latent	-1.097	0.249	-0.002	0.251	93.2	32.5
	Univariate	-1.1021	0.5044	0.0021	-0.0044	94.1	95.8
	Multivariate	-1.1035	0.5086	0.0035	-0.0086	95.6	92.4
	Shared	-1.102	0.505	0.002	-5.00E-03	94.1	95.8
n=500	Marginal	-1.102	0.505	0.002	-5.00E-03	94.1	95.8
	Latent	-1.102	0.253	0.002	0.248	94.3	2.3
	Univariate	-1.099	0.499	0.0009	0.0008	94.8	95
	Multivariate	-1.099	0.4999	-0.001	0.0001	94.7	95.4
	Shared	-1.099	0.499	-1.00E-03	0.001	94.9	95.1
n=1000	Marginal	-1.099	0.499	-1.00E-03	0.001	94.9	95.1
	Latent	-1.099	0.250	-0.001	0.251	95.1	0.0

Table 3: Estimates, bias and 95% Coverage for estimated effects for three binary outcomes with same covariates. β_0=1.0, β_1=-0.2

Sample size	Model	$\widehat{\beta}_0$	$\widehat{\beta}_1$	Bias ($\widehat{\beta}_0$)	Bias ($\widehat{\beta}_1$)	95% Cov. ($\widehat{\beta}_0$)	95% Cov. ($\widehat{\beta}_1$)
ρ**=0**							
	Univariate	1.711	-0.3756	-0.711	0.176	92.7	93.8
	Multivariate	0.973	-0.233	-0.053	-0.018	94.5	88.4
	Shared	1.011	-0.199	-0.011	-0.002	93.7	94.2
n=200	Marginal	0.997	-0.196	0.003	-0.004	95.7	94.3
	Latent	0.998	-0.091	0.002	-0.109	94.3	90.2
	Univariate	1.668	-0.341	-0.668	0.1414	89.2	88.5
	Multivariate	1.059	-0.187	-0.053	0.070	91.0	88.3
	Shared	1.012	-0.203	-0.012	0.0025	91.3	88.9
n=500	Marginal	1.003	-0.201	-0.0027	0.0008	98.5	95.6
	Latent	1.003	0.045	-0.003	-0.244	98.3	93.7
	Univariate	1.669	-0.355	-0.669	0.155	83.6	85.1
	Multivariate	1.059	-0.187	-0.053	0.070	91.0	90.4
	Shared	1.007	-0.201	-7.00E-03	0.0008	93.8	95.7
n=1000	Marginal	1.001	-0.200	5.00E-04	2.00E-04	95.3	97.8
	Latent	1.001	0.053	0.000	-0.253	97.7	92.9
$\rho = 0.3$							
	Univariate	1.699	-0.371	1-0.699	0.171	92.7	91.5
	Multivariate	1.037	-0.251	-0.037	0.051	89.8	92.5
	Shared	1.224	-0.248	-0.224	0.048	91.9	91.8
n=200	Marginal	1.011	-0.207	-0.011	7.00E-03	93.7	92.4
	Latent	1.026	0.237	-0.026	-0.437	98.4	93.4
	Univariate	1.688	-0.363	-0.688	0.1626	90.4	90.8
	Multivariate	1.040	-0.178	-0.040	0.0216	92.4	82.3
	Shared	1.213	-0.237	-0.213	0.037	88.5	89.8
n=500	Marginal	1.000	-0.199	0.000	-1.00E-03	98.5	93.6
	Latent	1.142	-0.115	-0.142	-0.085	94	93
	Univariate	1.674	-0.35	1-0.674	0.151	84.8	83
	Multivariate						
	Shared	1.211	-0.243	-0.211	0.043	80	78.7
n=1000	Marginal	1.000	-0.204	1.00E-04	0.004	99.5	92.1
	Latent	1.876	-0.177	-0.876	-0.033	92	94.2
$\rho = 0.8$							
	Univariate	1.702	-0.378	-0.702	0.1776	93.5	93.1
	Multivariate	1.027	-0.163	-0.0267	-0.0374	90.5	83.2
	Shared	3.917	-0.522	-2.917	0.322	88.2	91.1
n=200	Marginal	1.020	-0.212	-0.020	0.012	94.1	93.7
	Latent	1.717	0.631	-0.717	-0.831	90.6	83.4
	Univariate	1.669	-0.338	-0.669	0.138	90.2	89.4
	Multivariate	1.006	-0.212	-0.0063	0.0123	95.3	92.4
	Shared	3.674	-0.527	-2.674	3.27E-01	84.7	82.9
n=500	Marginal	1.003	-0.203	-0.004	3.00E-03	93.6	90.5
	Latent	1.657	0.703	-0.657	-0.903	89.2	10.8
	Univariate	1.623	-0.356	-0.6725	0.1563	83.8	84.4
	Multivariate	-	-	-	-	-	-
	Shared	3.416	-0.560	-2.42E	0.359	66.2	67.2
n=1000	Marginal	1.000	-0.204	0.0001	0.004	96.7	95.2
	Latent	1.642	0.720	-0.642	-0.920	91.8	86

B R source code for Chapter 5

```
#######################################Data arrangement
   ####################

#Reshape data
data_long <- melt(data, id.vars = c("id","wave","age_cont",
    "edn_yr","religion", "region", "incentive", "gen","
   aware"),
measure.vars = c("part","con"), variable.name="category",
   value.name="Y")

data2<-within(data_long, {
D1<-as.integer(category =="con")
D2<-as.integer(category== "part")
                                  })



#Stabilized weights for condom use
sw1 <- ipwtm(exposure = aware,
      family = "binomial",
      link = "logit",
      numerator = ~ factor(gen) +factor(wealth_cat) +factor
         (region) + factor(religion) + factor(incentive) +
         factor(perceived_part_hiv) +factor(marr)+factor(
         aids_info),
      denominator = ~factor(gen) +factor(wealth_cat) +
         factor(region) + factor(religion) + factor(
         incentive)+edn_yr +factor(perceived_part_hiv) +
```

```
         factor(marr) +factor(aids_info),
      id = id,
      timevar=wave,
      type="first",
      data = data2)



#Create censoring weights
lw <- ipwtm(exposure = ltfu_new,
      family = "binomial",
      link = "logit",
      numerator = ~ factor(aware)+factor(gen) +factor(
         wealth_cat)+factor(perceived_part_hiv)
      +factor(region) + factor(religion) + factor(incentive
         ) +factor(marr)+factor(aids_info),
      denominator = ~factor(aware)+factor(gen) +factor(
         wealth_cat)+factor(perceived_part_hiv)
      +factor(region) + factor(religion) +edn_yr +factor(
         marr) +factor(aids_info),
      id = id,
      timevar=wave,
      type="first",
      data = data2)

      lw<-lw$ipw.weights

#Create final weight
```

```
fwt = sw1*lw
data2 <-cbind(data2, fwt)

#Fit model using JGEE, with a bit of modification
#in the source file to incorporate the weights
#Variables of interest
vars <- c("age_cont", "edn_yr", "wealth_cat", "religion", "
   region",
            "perceived_part_hiv", "incentive", "distvct", "
               gender", "aids_info", "like_hiv")

#outcome model formula
outcome_frm <- as.formula(paste("Y~0+D1:factor(aware)+D2:
   factor(aware)␣+␣", paste(vars, collapse = "+"))
                                )
outcome_mod<-JGee1(formula=outcome_frm,id=id,data=data2,nr
   =2,na.action=NULL, family=binomial(link="logit"),
   corstr1="AR-1", Mv=NULL, corstr2="unstructured", beta_
   int=NULL, R1=NULL, R2=NULL, scale.fix=FALSE, scale.value
   =1, maxiter=30, tol=10^-3, silent=FALSE, weights =fwt)
```

Bibliography

Abadie, A., Drukker, D., Herr, J. L., & Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in stata. *The Stata Journal*, *4*(3), 290–311.

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, *74*(1), 235–267.

Adaralegbe, A., Adeoti, O., Ogbonda, S., Egeolu, O., & Ukoha, N. (2021). Sexual behaviors and hiv status: A population-based study among youths and adults in tanzania. *J Infect Dis Epidemiol*, *7*(8), 224. http://doi.org/10.23937/2474-3658/1510224

Adedimeji, A. A., Hoover, D. R., Shi, Q., Gard, T., Mutimura, E., Sinayobye, J. d., ... Anastos, K. (2015). Sexual behavior and risk practices of hiv positive and hiv negative rwandan women. *AIDS and behavior*, *19*(7), 1366–1378.

Agogo, G. O., Murphy, T. E., McAvay, G. J., & Allore, H. G. (2019). Joint modeling of concurrent binary outcomes in a longitudinal observational study using inverse probability of treatment weighting for treatment effect estimation. *Annals of epidemiology*, *35*(1), 53–58. https://doi.org/10.1016/j.annepidem.2019.04 .008

Anderson, T. W., & Olkin, I. (1985). Maximum-likelihood estimation of the parameters of a multivariate normal distribution. *Linear algebra and its applications*, *70*(1), 147–171. https://doi.org/10.1016/0024-3795(85)90049-7

Anglewicz, P. (2012). Migration, marital change, and hiv infection in malawi. *Demog-*

raphy, *49*(1), 239–265.

Anglewicz, P., & Clark, S. (2013). The effect of marriage and hiv risks on condom use acceptability in rural malawi. *Social Science & Medicine*, *97*(1), 29–40. https://doi.org/10.1016/j.socscimed.2013.06.024

Anglewicz, P., & Reniers, G. (2014). Hiv status, gender, and marriage dynamics among adults in rural malawi. *Studies in family planning*, *45*(4), 415–428.

Arvanitoyannis, I., Katsota, M., Psarra, E., Soufleros, E., & Kallithraka, S. (1999). Application of quality control methods for assessing wine authenticity: Use of multivariate analysis (chemometrics). *Trends in Food Science & Technology*, *10*(10), 321–336.

Austin, P. C. (2008). A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statistics in Medicine*, *27*(12), 2037–2049.

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effect of Confounding in Observational. *Multivariate Behavioural Research*.

Austin, P. C. (2014). A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*, *33*(6), 1057–1069.

Austin, P. C. (2018). Assessing the performance of the generalized propensity score for estimating the effect of quantitative or continuous exposures on binary outcomes. *Statistics in Medicine*, *37*(11), 1874–1894.

Austin, P. C., & Mamdani, M. M. (2006). A comparison of propensity score methods: a case-study estimating the effectiveness of post-ami statin use. *Statistics in Medicine*, *25*(12), 2084–2106.

Austin, P. C., & Stuart, E. A. (2015). Moving Towards Best Practice when Using Inverse Probability of Treatment Weighting (IPTW) using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies. *Statistics in Medicine*, *34*(28), 3661-3679.

Austin, P. C., & Stuart, E. A. (2017). The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Statistical Methods in Medical Research*, *26*(4), 1654–1670.

Bai, H., Zhong, Y., Gao, X., & Xu, W. (2020). Multivariate mixed response model with pairwise composite-likelihood method. *Stats*, *3*(3), 203–220.

Baranov, V., & Kohler, H.-P. (2018). The impact of aids treatment on savings and human capital investment in malawi. *American Economic Journal: Applied Economics*, *10*(1), 266–306.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R package version*, *2*(1), 74.

Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, *2*(4), 358–377.

Bodnar, L. M., Davidian, M., Siega-Riz, A. M., & Tsiatis, A. A. (2004). Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology. *American Journal of Epidemiology*, *159*(10), 926–934.

Bonat, W. H. (2018). Multiple response variables regression models in r: the mcglm package. *Journal of Statistical Software*, *84*(1), 1–30.

Brobbey, A., Wiebe, S., Nettel-Aguirre, A., Josephson, C. B., Williamson, T., Lix, L. M., & Sajobi, T. T. (2021). Repeated measures discriminant analysis using multivariate generalized estimation equations. *Statistical Methods in Medical Research*, *31*(4), 646–657. http://doi.org/10.1177/09622802211032705

Brumback, B. A., He, Z., Prasad, M., Freeman, M. C., & Rheingans, R. (2014). Using structural-nested models to estimate the effect of cluster-level adherence on individual-level outcomes with a three-armed cluster-randomized trial. *Statistics in medicine*, *33*(9), 1490–1502.

Bunnell, R., Opio, A., Musinguzi, J., Kirungi, W., Ekwaru, P., Mishra, V., . . . Mermin, J. (2008). Hiv transmission risk behavior among hiv-infected adults in uganda: results of a nationally representative survey. *Aids*, *22*(5), 617–624.

Cao, W., Tsiatis, A. A., & Davidian, M. (2009). Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika*, *96*(3), 723–734.

Carey, V., Zeger, S. L., & Diggle, P. (1993). Modelling multivariate binary data with alternating logistic regressions. *Biometrika*, *80*(3), 517–526.

Chesnaye, N. C., Stel, V. S., Tripepi, G., Dekker, F. W., Fu, E. L., Zoccali, C., & Jager, K. J. (2021). An introduction to inverse probability of treatment weighting in observational research. *Clinical Kidney Journal*, *15*(1), 14-20. doi: 10.1093/ckj/sfab158

Chinkonde, J. R., Hem, M. H., & Sundby, J. (2012). Hiv and infant feeding in malawi: public health simplicity in complex social and cultural contexts. *BMC Public Health*, *12*(1), 1–9.

Chintalapudi, N., Hamela, G., Mofolo, I., Maman, S., Hosseinipour, M. C., Hoffman, I. F., & Flax, V. L. (2018). Infant and young child feeding decision making and practices: Malawian mothers' and fathers' roles in the context of hiv. *Journal of Human Lactation*, *34*(1), 68–76.

Cho, H. (2016). The analysis of multivariate longitudinal data using multivariate marginal models. *Journal of Multivariate Analysis*, *143*(1), 481–491. https://doi.org/10.1016/j.jmva.2015.10.012

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, *24*(2), 295–313. https://doi.org/10.2307/2528036

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology*, *168*(6), 656–664.

Cole, S. R., & Hernan, M. A. (2008). An Introduction to Propensity Score Methods for Reducing the Effect of Confounding in Observational. *American Journal of Epidemiology*, *46*(3), 399–424. http://doi.org/10.1080/00273171.2011.568786

Cole, S. R., Hernán, M. A., Robins, J. M., Anastos, K., Chmiel, J., Detels, R., ... others (2003). Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *American Journal of Epidemiology*, *158*(7), 687–694.

Cui, Y., Michael, H., Tanser, F., & Tchetgen Tchetgen, E. (2023). Instrumental variable estimation of the marginal structural cox model for time-varying treatments. *Biometrika*, *110*(1), 101–118.

d'Agostino, R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in medicine*, *17*(19), 2265–2281.

Danaei, G., Rodríguez, L. A. G., Cantero, O. F., Logan, R., & Hernán, M. A. (2013). Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical methods in medical research*, *22*(1), 70–96.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151–161.

De Onis, M., Blossner, M., Organization, W. H., et al. (1997). *Who global database on*

child growth and malnutrition (Tech. Rep.). World Health Organization.

Dessau, R. B., & Pipper, C. B. (2008). "r"–project for statistical computing. *Ugeskrift for laeger*, *170*(5), 328–330.

Dewey, K. (2003). *Guiding principles for complementary feeding of the breastfed child*. Pan American Health Organization/WHO.

Dreyfuss, M. L., Msamanga, G. I., Spiegelman, D., Hunter, D. J., Urassa, E. J., Hertzmark, E., & Fawzi, W. W. (2001). Determinants of low birth weight among HIV-infected pregnant women in Tanzania. *The American Journal of Clinical Nutrition*, *74*(6), 814–26. http://www.ncbi.nlm.nih.gov/pubmed/11722965

Dukes, O., & Vansteelandt, S. (2018). A note on g-estimation of causal risk ratios. *American Journal of Epidemiology*, *187*(5), 1079–1084.

Egbewale, B. E., Lewis, M., & Sim, J. (2014). Bias, precision and statistical power of analysis of covariance in the analysis of randomized trials with baseline imbalance: a simulation study. *BMC medical research methodology*, *14*(1), 1–12.

Egger, M., Coleman, M., Ward, J., Reading, J., & Williams, H. (1985). Uses and abuses of analysis of covariance in clinical trials. *Controlled Clinical Trials*, *6*(1), 12–24.

Fedor, T. M., Kohler, H.-P., & Behrman, J. R. (2015). The impact of married individuals learning hiv status in malawi: divorce, number of sexual partners, and condom use with spouses. *Demography*, *52*(1), 259–280.

Feng, P., Zhou, X.-H., Zou, Q.-M., Fan, M.-Y., & Li, X.-S. (2012). Generalized propensity score for estimating the average treatment effect of multiple treatments. *Statistics in medicine*, *31*(7), 681–697.

Fieuws, S., & Verbeke, G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics*, *62*(2), 424–431.

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian,

M. (2011). Doubly robust estimation of causal effects. *American journal of epidemiology*, *173*(7), 761–767.

Gao, Q., Zhang, Y., Liang, J., Sun, H., & Wang, T. (2021). High-dimensional generalized propensity score with application to omics data. *Brief Bioinformatics*, *22*(6), bbab331. http://doi.org/10.1093/bib/bbab331

George, G., Beckett, S., Cawood, C., Khanyile, D., Govender, K., & Kharsany, A. B. (2019). Impact of hiv testing and treatment services on risky sexual behaviour in the umgungundlovu district, kwazulu-natal, south africa: a cross-sectional study. *AIDS research and therapy*, *16*(1), 1–11.

Government of Malawi, N., & ICF. (2017). *Malawi demographic and health survey 2015-16 report* (Tech. Rep.). National Statistics Office Malawi and ICF.

Guiella, G., & Madise, N. J. (2007). Hiv/aids and sexual-risk behaviors among adolescents: Factors influencing the use of condoms in burkina faso. *African journal of reproductive health*, *11*(3), 182–196.

Guo, S., Barth, R. P., & Gibbons, C. (2006). Propensity score matching strategies for evaluating substance abuse services for child welfare clients. *Children and Youth Services Review*, *28*(4), 357–383.

Hernan, M. A., Brumback, B., & Robins, J. (2000). *Marginal Structural Models to Estimate the causal effect of zidovudine on the survival of hiv-positive men.* LWW.

Hernan, M. A., Brumback, B. A., & Robins, J. M. (2002). Estimating the causal effect of zidovudine on cd4 count with a marginal structural model for repeated measures. *Statistics in medicine*, *21*(12), 1689–1709.

Hernán, M. A., & Robins, J. M. (2020). *Causal inference. what if?* Havard University.

Hirano, K., & Imbens, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes research methodology*, *2*(3-4), 259–278.

Højsgaard, S., Halekoh, U., & Yan, J. (2006). The r package geepack for generalized estimating equations. *Journal of statistical software*, *15*(1), 1–11.

Holland, D. (1986). Causal Inference Using Potential Outcomes: Design, Modeling, Decision. *American Statistical Association*.

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, *47*(260), 663–685.

Howe, C. J., Cole, S. R., Mehta, S. H., & Kirk, G. D. (2012). Estimating the effects of multiple time-varying exposures using joint marginal structural models: alcohol consumption, injection drug use, and hiv acquisition. *Epidemiology (Cambridge, Mass.)*, *23*(4), 574–582.

Huerga, H., Venables, E., Ben-Farhat, J., van Cutsem, G., Ellman, T., & Kenyon, C. (2017). Higher risk sexual behaviour is associated with unawareness of hiv-positivity and lack of viral suppression–implications for treatment as prevention. *Scientific Reports*, *7*(1), 1–7.

Imai, K., & Ratkovic, M. (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association*, *110*(511), 1013–1023.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, *87*(3), 706–710.

Inan, G. (2015). Joint generalized estimating equation solver (jgee). *R package version*. https://cran.r-project.org/web//packages/JGEE/JGEE.pdf

Inan, G., & Yucel, R. (2017). Joint gees for multivariate correlated data with incomplete binary outcomes. *Journal of Applied Statistics*, *44*(11), 1920–1937.

Jager, K., Zoccali, C., Macleod, A., & Dekker, F. (2008). Confounding: what it is and how to deal with it. *Kidney international*, *73*(3), 256–260.

Jewkes, R., Dunkle, K., Nduna, M., Levin, J., Jama, N., Khuzwayo, N., ... Duvvury, N. (2006). Factors associated with HIV sero-status in young rural South African women: connections between intimate partner violence and HIV. *International Journal of Epidemiology*, *35*(6), 1461–1468. http://doi.org/10.1093/ije/dyl218

Kahlert, J., Gribsholt, S. B., Gammelager, H., Dekkers, O. M., & Luta, G. (2017). Control of confounding in the analysis phase–an overview for clinicians. *Clinical Epidemiology*, *9*(1), 195–204. http://doi.org/10.2147/CLEP.S129886

Kalia, S., Saarela, O., Escobar, M., Moineddin, R., & Greiver, M. (2023). Estimation of marginal structural models under irregular visits and unmeasured confounder: calibrated inverse probability weights. *BMC Medical Research Methodology*, *23*(1), 1–16.

Kamenju, P., Liu, E., Hertzmark, E., Spiegelman, D., Kisenge, R., Kupka, R., ... Fawzi, W. W. (2017). Nutritional status and complementary feeding among hiv-exposed infants: a prospective cohort study. *Maternal & Child Nutrition*, *13*(3), e12358.

Karim, M. E., Platt, R. W., & Group, B. S. (2017). Estimating inverse probability weights using super learner when weight-model specification is unknown in a marginal structural cox model context. *Statistics in Medicine*, *36*(13), 2032–2047.

Kassa, T., Meshesha, B., Haji, Y., & Ebrahim, J. (2016). Appropriate complementary feeding practices and associated factors among mothers of children age 6–23 months in southern ethiopia, 2015. *BMC Pediatrics*, *16*(1), 1–10.

Kendall, J. (2003). Designing a research project: randomised controlled trials and their principles. *Emergency medicine journal: EMJ*, *20*(2), 164.

Kennedy, E. H., Kangovi, S., & Mitra, N. (2019). Estimating scaled treatment effects with multiple outcomes. *Statistical Methods in Medical Research*, *28*(4), 1094–1104.

Keogh, R. H., Daniel, R. M., Vanderweele, T. J., & Vansteelandt, S. (2017). Analysis of longitudinal studies with repeated outcome measures: adjusting for time-dependent confounding using conventional methods. *American Journal of Epidemiology*, *187*(5), 1085–1092.

Kohler, H.-P., Watkins, S. C., Behrman, J. R., Anglewicz, P., Kohler, I. V., Thornton, R. L., ... Kalilani-Phiri, L. (2014). Cohort profile: The malawi longitudinal study of families and health (mlsfh). *International Journal of Epidemiology*, *44*(2), 394–404.

Kohler, I., Bandawe, C., Ciancio, A., Kämpfen, F., Payne, C. F., Mwera, J., ... Kohler, H.-P. (2020). Cohort profile: the mature adults cohort of the malawi longitudinal study of families and health (mlsfh-mac). *BMJ open*, *10*(10), e038232.

Kumar, D., Goel, N., Mittal, P. C., & Misra, P. (2006). Influence of infant-feeding practices on nutritional status of under-five children. *The Indian Journal of Pediatrics*, *73*(5), 417–421.

Kurth, T., Walker, A. M., Glynn, R. J., Chan, K. A., Gaziano, J. M., Berger, K., & Robins, J. M. (2006). Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *American Journal of Epidemiology*, *163*(3), 262–270.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*(4), 963–974.

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*(3), 337–346.

Leisch, F., Weingessel, A., & Hornik, K. (1998). On the generation of correlated artificial binary data. *WU Working Paper Series*.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Liang, K.-Y., Zeger, S. L., & Qaqish, B. (1992). Multivariate regression analyses for categorical data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *54*(1), 3–24.

Lipkovich, I., Mallinckrodt, C. H., & Faries, D. E. (2012). The challenges of evaluating dose response in flexible-dose trials using marginal structural models. *Pharmaceutical Statistics*, *11*(6), 485–493.

Lipsitz, S. R., Laird, N. M., & Harrington, D. P. (1991). Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. *Biometrika*, *78*(1), 153–160.

Lu, F., Xue, L., & Hu, Y. (2020). Robust estimation for the correlation matrix of multivariate longitudinal data. *Journal of Statistical Computation and Simulation*, *90*(13), 2473–2496.

Lunceford, J. K., & Davidian, M. (2004). Stratifcation and Weighting via the Propensity Score in Estimation of Causal Treatment Effects: a Comparative Study. *Statistics in Medicine*, *23*(19), 2937-2960. https://doi.org/10.1002/sim.1903

Lupparelli, M., & Mattei, A. (2017). *Causal inference for binary non-independent outcomes*. *Pre-print*. https://doi.org/10.48550/arXiv.1710.07039

Lusivika-Nzinga, C., Selinger-Leneman, H., Grabar, S., Costagliola, D., & Carrat, F. (2017). Performance of the marginal structural cox model for estimating individual and joined effects of treatments given in combination. *BMC Medical Research Methodology*, *17*(1), 1–11.

Machira, K., Maonga, B., & Chirwa, T. (2020). Determinants of risky sexual behavioral practices among teen-girls in malawi. *The Open Public Health Journal*, *13*(1), 839–849. http://doi.org/10.2174/1874944502013010839

Mathur, M. B., Ding, P., Riddell, C. A., & VanderWeele, T. J. (2018). Website and r package for computing e-values. *Epidemiology (Cambridge, Mass.)*, *29*(5),

e45–e47.

Mattei, A., Li, F., Mealli, F., et al. (2013). Exploiting multiple outcomes in bayesian principal stratification analysis with application to the evaluation of a job training program. *The Annals of Applied Statistics*, *7*(4), 2336–2360.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, *32*(19), 3388–3414.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004, dec). Propensity Score Estimation With Boosted Regression for Evaluating Causal Effects in Observational Studies. *Psychological Methods*, *9*(4), 403–425. doi: 10.1037/1082-989X.9.4.403

McConnell, K. J., & Lindner, S. (2019). Estimating treatment effects with machine learning. *Health services research*, *54*(6), 1273–1282.

MDHS. (2010). *Malawi demographic and health survey 2010 report* (Tech. Rep.). National Statistics Office Malawi.

Michael, H., Cui, Y., Lorch, S., & Tchetgen, E. T. (2020). Instrumental variable estimation of marginal structural mean models for time-varying treatment. *Preprint*. https://doi.org/10.48550/arXiv.2004.11769

Moodie, E. E., Stephens, D. A., & Klein, M. B. (2014). A marginal structural model for multiple-outcome survival data: assessing the impact of injection drug use on several causes of death in the canadian co-infection cohort. *Statistics in Medicine*, *33*(8), 1409–1425.

Naik, D. N., & Rao, S. S. (2001). Analysis of multivariate repeated measures data with a kronecker product structured covariance matrix. *Journal of Applied Statistics*, *28*(1), 91–105.

Naimi, A. I., Cole, S. R., & Kennedy, E. H. (2017). An introduction to g methods.

International journal of epidemiology, *46*(2), 756–762.

Oliveira, R., & Teixeira-Pinto, A. (2015). Analyzing multiple outcomes: Is it really worth the use of multivariate linear regression? *Journal of Biometrics & Biostatistics*, *6*(4). http:\\doi.org\10.4172/2155-6180.1000256

Ortblad, K. F., Musoke, D. K., Ngabirano, T., Salomon, J. A., Haberer, J. E., McConnell, M., . . . Bärnighausen, T. (2019). Is knowledge of hiv status associated with sexual behaviours? a fixed effects analysis of a female sex worker cohort in urban uganda. *Journal of the International AIDS Society*, *22*(7), e25336.

Palk, L., & Blower, S. (2018). Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the hiv epidemic in malawi. *BMC Medicine*, *16*(1), 1–9.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, *57*(1), 120–125.

Pan, W., & Bai, H. (2015). Propensity score interval matching: using bootstrap confidence intervals for accommodating estimation errors of propensity scores. *BMC medical research methodology*, *15*(1), 1–9.

Pega, F., Blakely, T., Glymour, M. M., Carter, K. N., & Kawachi, I. (2016). Using marginal structural modeling to estimate the cumulative impact of an unconditional tax credit on self-rated health. *American Journal of Epidemiology*, *183*(4), 315–324.

Picciotto, S., & Neophytou, A. M. (2016). G-estimation of structural nested models: recent applications in two subfields of epidemiology. *Current Epidemiology Reports*, *3*(3), 242–251.

Platt, R. W., Brookhart, M. A., Cole, S. R., Westreich, D., & Schisterman, E. F. (2013). An information criterion for marginal structural models. *Statistics in Medicine*, *32*(8), 1383–1393.

Porter, L., Hao, L., Bishai, D., Serwadda, D., Wawer, M. J., Lutalo, T., & Gray, R. (2004). Hiv status and union dissolution in sub-saharan africa: the case of rakai, uganda. *Demography*, *41*(3), 465–482.

Pourhoseingholi, M. A., Baghestani, A. R., & Vahedi, M. (2012). How to control confounding effects by statistical analysis. *Gastroenterology and hepatology from bed to bench*, *5*(2), 79–83.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). Gllamm manual. *Working Paper series*.

Reniers, G., & Armbruster, B. (2012). Hiv status awareness, partnership dissolution and hiv transmission in generalized epidemics. *PLoS One*, *7*(12), e50669.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling*, *7*(9-12), 1393–1512.

Robins, J. M. (1997). Causal inference from complex longitudinal data. In *Latent variable modeling and applications to causality* (pp. 69–117). Springer.

Robins, J. M. (1999). Association, causation, and marginal structural models. *Synthese*, *121*(1-2), 151–179.

Robins, J. M. (2000). Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. In *Statistical models in epidemiology, the environment, and clinical trials* (pp. 95–133). Springer.

Robins, J. M., Greenland, S., & Hu, F.-C. (1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of the American Statistical Association*, *94*(447), 687–700.

Robins, J. M., & Hernan, M. A. (2009). Estimation of the causal effects of time varying exposures. In *Longitudinal data analysis* (pp. 553–99). Chapman and Hall/CRC.

Robins, J. M., Hernan, M. A., & Brumback, B. (2000). *Marginal structural models and causal inference in epidemiology.* LWW.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, *89*(427), 846–866.

Rollins, N. C., Coovadia, H. M., Bland, R. M., Coutsoudis, A., Bennish, M. L., Patel, D., & Newell, M.-L. (2007). Pregnancy outcomes in hiv-infected and uninfected women in rural and urban south africa. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *44*(3), 321–328.

Rosenbaum, P. R. (2002). Observational studies. In *Observational studies* (pp. 1–17). Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*, *79*(387), 516–524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Rosenberg, M. S., Gómez-Olivé, F. X., Rohr, J. K., Houle, B. C., Kabudula, C. W., Wagner, R. G., . . . others (2017). Sexual behaviors and hiv status: a population-based study among older adults in rural south africa. *Journal of acquired immune deficiency syndromes (1999)*, *74*(1), e9.

Roy, A., & Khattree, R. (2005). On implementation of a test for kronecker product covariance structure for multivariate repeated measures data. *Statistical Methodology*, *2*(4), 297–306.

Roy, J., Lin, X., & Ryan, L. M. (2003). Scaled marginal models for multiple continuous outcomes. *Biostatistics*, *4*(3), 371–383.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, *66*(5), 688.

Rubin, D. B. (2006). Matching to remove bias in observational studies. In *Matched sampling for causal effects* (pp. 62–80). Cambridge University Press. http://doi.org/10.1017/CBO9780511810725.007

Sainani, K. L. (2018). *Instrumental variables: uses and limitations* (Vol. 10) (No. 3). Elsevier.

Saint-Mont, U. (2015). Randomization does not help much, comparability does. *PloS one*, *10*(7), e0132102. https://doi.org/10.1371/journal.pone.0132102

Sammel, M., Lin, X., & Ryan, L. (1999). Multivariate linear mixed models for multiple outcomes. *Statistics in medicine*, *18*(17-18), 2479–2492.

Sammel, M., & Ryan, L. M. (1996). Latent variable models with fixed effects. *Biometrics*, *52*(2), 650–663.

Sammel, M., Ryan, L. M., & Legler, J. M. (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *59*(3), 667–678.

Senn, S. (1989). Covariate imbalance and random allocation in clinical trials. *Statistics in Medicine*, *8*(4), 467–475.

Srivastava, M. S., von Rosen, T., & Von Rosen, D. (2008). Models with a kronecker product covariance structure: estimation and testing. *Mathematical Methods of Statistics*, *17*(4), 357–370.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical*

Statistics, *25*(1), 1–21. http://doi.org/10.1214/09-STS313

Suarez, D., Borràs, R., & Basagaña, X. (2011). Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. *Epidemiology*, *22*(4), 586–588.

Sullivan Pepe, M., & Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics-Simulation and Computation*, *23*(4), 939–951.

Suresh, K. (2011). An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences*, *4*(1), 8.

Swan, T. (2006). *Generalized estimating equations when the response variable has a tweedie distribution: An application for multi-site rainfall modelling* (Unpublished doctoral dissertation). University of Southern Queensland.

Tager, I. B., Haight, T., Sternfeld, B., Yu, Z., & van Der Laan, M. (2004). Effects of physical activity and body composition on functional limitation in the elderly: application of the marginal structural model. *Epidemiology*, *15*(4), 479–493.

Tan, Z. (2010a). Bounded, efficient and doubly robust estimation with inverse weighting. *Biometrika*, *97*(3), 661–682.

Tan, Z. (2010b). Marginal and nested structural models using instrumental variables. *Journal of the American Statistical Association*, *105*(489), 157–169.

Taubman, S. L., Robins, J. M., Mittleman, M. A., & Hernán, M. A. (2009). Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *International Journal of Epidemiology*, *38*(6), 1599–1611.

Teixeira-Pinto, A., & Mauri, L. (2011). Statistical analysis of noncommensurate multiple outcomes. *Circulation: Cardiovascular Quality and Outcomes*, *4*(6), 650–656.

Tsai, A. C., Weiser, S. D., Petersen, M. L., Ragland, K., Kushel, M. B., & Bangsberg, D. R. (2010). A marginal structural model to estimate the causal effect of antidepressant medication treatment on viral suppression among homeless and marginally housed persons with hiv. *Archives of General Psychiatry*, *67*(12), 1282–1290.

Tu, C., Koh, W. Y., & Jiao, S. (2013). Using generalized doubly robust estimator to estimate average treatment effects of multiple treatments in observational studies. *Journal of Statistical Computation and Simulation*, *83*(8), 1518–1526.

Twabi, H. S., Manda, S. O., & Small, D. S. (2020). Assessing the effects of maternal hiv infection on pregnancy outcomes using cross-sectional data in malawi. *BMC Public Health*, *20*(1), 1–15.

Twabi, H. S., Manda, S. O., & Small, D. S. (2021). Evaluating the effect of appropriate complementary feeding practices on child growth in malawi using cross-sectional data: An application of propensity score matching. *Frontiers in Nutrition*, *8*(714232). http://doi.org/10.3389/fnut.2021.714232

Twabi, H. S., Manda, S. O., Small, D. S., & Kohler, H.-P. (2022). A marginal structural model for estimation of the effect of hiv positivity awareness on risky sexual behavior. *Biostatistics & Epidemiology*, *6*(2), 309–326.

UNAIDS. (2015). Understanding fast-track: Accelerating action to end the aids epidemic by 2030. *Author*.

Vaahtera, M., Kulmala, T., Hietanen, A., Ndekha, M., Cullinan, T., Salin, M.-L., & Ashorn, P. (2001). Breastfeeding and complementary feeding practices in rural malawi. *Acta Paediatrica*, *90*(3), 328–332.

van der Wal, W. M., Geskus, R. B., et al. (2011). ipw: an r package for inverse probability weighting. *J Stat Softw*, *43*(13), 1–23.

VanderWeele, T. J. (2009). Marginal structural models for the estimation of direct and

indirect effects. *Epidemiology*, *20*(1), 18–26.

VanderWeele, T. J. (2017). Outcome-wide epidemiology. *Epidemiology (Cambridge, Mass.)*, *28*(3), 399.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: introducing the e-value. *Annals of Internal Medicine*, *167*(4), 268–274.

VanderWeele, T. J., Jackson, J. W., & Li, S. (2016). Causal inference and longitudinal data: a case study of religion and mental health. *Social Psychiatry and Psychiatric Epidemiology*, *51*(11), 1457–1466.

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: a new template for empirical studies. *Statistical Science*, *35*(3), 437–466.

Vansteelandt, S., Joffe, M., et al. (2014). Structural nested models and g-estimation: the partially realized promise. *Statistical Science*, *29*(4), 707–731.

Ventura-Filipe, E. M., & Newman, S. P. (1998). Influence of hiv positive status on sexual behavior among males. *Revista de Saúde Pública*, *32*(6), 503–513.

Verbeke, G., Fieuws, S., Molenberghs, G., & Davidian, M. (2014). The analysis of multivariate longitudinal data: a review. *Statistical Methods in Medical Research*, *23*(1), 42–59.

Walters, C. N., Rakotomanana, H., Komakech, J. J., & Stoecker, B. J. (2019). Maternal determinants of optimal breastfeeding and complementary feeding and their association with child undernutrition in malawi (2015–2016). *BMC Public Health*, *19*(1), 1503. http://doi.org/10.1186/s12889-019-7877-8

Werner, K., Jansson, M., & Stoica, P. (2008). On estimation of covariance matrices with kronecker product structure. *IEEE Transactions on Signal Processing*, *56*(2), 478–491.

WHO. (2006). *Who child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. Author.

WHO. (2009). The importance of infant and young child feeding and recommended practices. *Author*.

WHO. (2010). *Indicators for assessing infant and young child feeding practices: part 2: measurement*. Author.

WHO. (2017). *Global targets 2025 to improve maternal, infant and young children nutrition* (Tech. Rep.). Author.

WHO, & UNICEF. (2003). *Global strategy for infant and young child feeding*. Author

Winkens, B., van Breukelen, G. J., Schouten, H. J., & Berger, M. P. (2007). Randomized clinical trials with a pre-and a post-treatment measurement: repeated measures versus ancova models. *Contemporary clinical trials*, *28*(6), 713–719.

Wondemagegn, F., & Berkessa, T. (2020). High level risky sexual behavior among persons living with hiv in the urban setting of the highest hiv prevalent areas in ethiopia: Implications for interventions. *PloS one*, *15*(11), e0242701.

Yang, S., Eaton, C. B., McAlindon, T. E., & Lapane, K. L. (2015). Effects of glucosamine and chondroitin supplementation on knee osteoarthritis: an analysis with marginal structural models. *Arthritis & Rheumatology*, *67*(3), 714–723.

Yu, J., & Wang, X. (2021). Maximum likelihood estimation for multivariate normal distribution with hierarchical missing data. *International Journal of Statistics and Applied Mathematics*, *6*(3), 16–19.

Zambia Statistics Agency, M. o. H. Z., & ICF. (2019). *Zambia demographic and health survey 2018* (Tech. Rep.). Author.

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and

continuous outcomes. *Biometrics*, *42*(1), 121–130.

Zeger, S. L., Liang, K.-Y., & Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, *44*(4), 1049–1060.

Zeng, L., & Cook, R. J. (2007). Transition models for multivariate longitudinal binary data. *Journal of the American Statistical Association*, *102*(477), 211–223.

Zhou, L., Lin, H., Song, X., & Li, Y. (2014). Selection of latent variables for multiple mixed-outcome models. *Scandinavian Journal of Statistics*, *41*(4), 1064–1082.

Zhu, Y., Ghosh, D., Mitra, N., & Mukherjee, B. (2014). A data-adaptive strategy for inverse weighted estimation of causal effects. *Health Services and Outcomes Research Methodology*, *14*(3), 69–91.

Sample Publications

Twabi *et al. BMC Public Health* (2020) 20:974
https://doi.org/10.1186/s12889-020-09046-0

BMC Public Health

RESEARCH ARTICLE Open Access

Assessing the effects of maternal HIV infection on pregnancy outcomes using cross-sectional data in Malawi

Halima S. Twabi[1*], Samuel O. Manda[2,3,4] and Dylan S. Small[5]

Abstract

Background: Several studies have shown that maternal HIV infection is associated with adverse pregnancy outcomes such as low birth weight and perinatal mortality. However, the association is conflicted with the effect of antiretroviral therapy (ART) on the pregnancy outcomes and it remains unexamined. If the association is confirmed then it would guide policy makers towards more effective prevention of mother to child HIV transmission interventions. Using methods for matching possible confounders, the objectives of the study were to assess the effect of maternal HIV infection on birth weight and perinatal mortality and to investigate the effect of ART on these two pregnancy outcomes in HIV-infected women.

Methods: Data on 4111 and 4759 children, born within five years of the 2010 and 2015-16 Malawi Demographic and Health Surveys (MDHS) respectively, whose mothers had an HIV test result, were analysed. A best balancing method was chosen from a set of covariate balance methods namely, the 1:1 nearest neighbour (NN) matching, matching on the propensity score (PS) and inverse weighting on the PS. HIV and ART data were only available in the MDHS 2010, permitting an assessment of the moderating effect of ART on the association between maternal HIV infection and birth weight and perinatal mortality.

Results: The overall average birth weight was 3227.9g (95% CI: 3206.4, 3249.5) in 2010 and 3226.4g (95%: 3205.6, 3247.2) in 2015-16 and perinatal mortality was 3.8% (95%: 3.2, 4.3) in 2010 and 3.5% (95%: 2.8, 3.8) in 2015-16. The prevalence of HIV among the mothers was 11.1% (95%: 10.1, 12.0) and 9.2% (95% CI: 8.4, 10.1) in 2010 and 2015-16, respectively. In 2010, maternal HIV infection was negatively associated with birth weight (mean= -25.3g, 95% CI:(-95.5, -7.4)) and in 2015-16 it was positively associated with birth weight (mean= 116.3g, 95% CI:(27.8, 204.7)). Perinatal mortality was higher in infants of HIV-infected mothers compared to infants of HIV-uninfected mothers (OR = 1.5, 95% CI:(1.1 - 3.1)) in 2010, while there was no difference in the rate in 2015-16 (OR = 1.0, 95% CI:(0.4, 1.6)). ART was not associated with birth weight, however, it was associated with perinatal mortality (OR=3.9, 95% CI:(1.1, 14.8)).

Conclusion: The study has found that maternal HIV infection had an adverse effect on birth weight and perinatal mortality in 2010. Birth weight was not dependent on ART uptake but perinatal mortality was higher among infants of HIV-infected mothers who were not on ART. The higher birth weight among HIV-infected mothers and similarity in perinatal mortality with HIV-uninfected mothers in 2015-16 may be indicative of successes of interventions within the PMTCT program in Malawi.

Keywords: Propensity score, Confounders, Maternal HIV, Inverse weighting, Perinatal mortality, Antiretroviral therapy

*Correspondence: htwabi@cc.ac.mw
[1]Department of Mathematical Sciences, University of Malawi, Zomba, Malawi
Full list of author information is available at the end of the article

Background

Maternal HIV infection has been shown to be associated with increased rates of adverse pregnancy outcomes such as low birth weight (LBW) (weighing less than 2500g at birth) and increased perinatal and early neonatal mortality [1–8]. The adverse impacts are largely due to maternal and obstetric complications among HIV-infected women [9, 10]. Nearly 20.5 million infants globally were born with low birth weight in the year 2015 [11]. Studies have shown that low birth is associated with high levels of child morbidly and mortality [12]. Infant mortality and morbidity persist in sub-Saharan African (SSA) region in which the HIV prevalence among pregnant women is high [13]. Identifying important and modifiable factors that influence LBW and perinatal mortality in a population with a high HIV burden, improves understanding of the pathways underlying the association between maternal HIV infection and pregnancy outcomes. In addition, it provides supporting scientific evidence for possible prevention and intervention within the Prevention of Mother-to-child HIV transmission (PMTCT) programme.

A growing number of studies have shown conflicting evidence on the association between maternal HIV and adverse pregnancy outcomes such as LBW. Some studies have shown that antiretroviral therapy (ART) among HIV-infected mothers is associated with adverse outcomes such as low birth weight (LBW), intrauterine growth restriction (IUGR), preterm delivery (PTD) and stillbirths [14–18]. On the other hand, several studies have found that sustained intake of ART among HIV-infected mothers reduces adverse pregnancy outcomes such as stillbirths, low birth weight and prematurity [16, 19]. Despite the disparity in the evidence, statistical methods can be used to determine the moderating effect of ART on the association between maternal HIV-infection and birth weight and perinatal mortality.

Most of the evidence on the association between maternal HIV and pregnancy outcomes comes from observational studies. However, establishing the impact of maternal HIV infection on birth weight and perinatal mortality is challenging because there could be other purported factors besides maternal HIV infection associated with pregnancy outcomes. The presence of confounding factors misrepresents the impact of maternal HIV infection on pregnancy outcomes as the distribution of these factors may be different between HIV-infected and HIV-uninfected mothers. Maternal anaemia, genitourinary tract infection in pregnancy, hypertension, history of birth weight and birth order are some of the factors known to be associated with pregnancy outcomes [14–17, 20–22]. Additionally, socio-economic and environmental factors such as maternal education, maternal age, poverty levels, air pollutants and smoking are associated with pregnancy outcomes [22–26]. To ensure comparability between HIV-infected and HIV-uninfected mothers, a randomised control trial (RCT) would ideally need to be conducted. However, since most HIV studies are observational, using methods developed by Rosenbaum and Rubin [27], one could achieve the same task of balancing confounder variables between HIV-infected and HIV-uninfected mothers. For example, one such method is based on Propensity Score (PS) matching, which aims to emulate the randomisation procedure of RCTs, as the distribution of measured confounders between exposure groups is made similar by using the PS [27–33, 33–35].

Ascertaining the association between maternal HIV infection and birth weight and perinatal mortality could be conflicted by the association between ART and the outcomes in HIV-infected mothers. The study therefore aimed at assessing the effect of maternal HIV infection on birth weight and perinatal mortality, and to investigate the effect of ART on the two pregnancy outcomes in HIV-infected women. This was done within an appropriate application of causal inference statistical methods to balance a set of confounder variables. The association between maternal HIV infection and birth weight and perinatal mortality with and without controlling for ART uptake in a high HIV epidemic and ART uptake population, remains largely untested. Confirmation of these associations would help policy makers obtain robust evidence-based findings that would assist in providing effective prevention of mother to child HIV transmission (PMTCT) interventions.

Methods

Study data

The data from this study were obtained from the 2010 and 2015-16 Malawi Demographic and Health Surveys (MDHS). These are nationally representative household surveys that provide data for a wide range of population, health, and nutrition indicators in low-income and middle-income (LIMICs). The Demographic and Health Surveys are implemented by the Demographic Health Survey Programme which is funded by the United States Agency for International Development (USAID). The two Malawian surveys provided up-to-date information on current HIV trends, child health and nutrition. Additionally, the 2010 survey provided information on ART. Parents of infants included in the surveys had signed an informed consent. Details of the sampling design for the MDHS can be obtained from the 2010 and 2015-16 reports [36, 37]. The study analyzed data on 4,111 and 4,759 children born to mothers who had an HIV result during the respective surveys. The outcome variables of interest generated were: Birth weight (grams) and perinatal mortality (0= Alive 1= Dead). The exposure variable of interest was mothers' HIV-infection status (0 = HIV

negative, 1 = HIV positive). The moderator variable was ART status among HIV-infected mothers (0 = Not on ART, 1 = On ART) using the 2010 MDHS. The effect estimates were obtained for the effect of maternal HIV infection on birth weight and perinatal mortality.

From the surveys, the overall average birth weight was 3227.9g (95% CI: 3206.4, 3249.5) and 3226.4g (95% CI: 3205.6, 3247.2) in 2010 and 2015-16, respectively. In 2010 the mean birth weight of children born to HIV-infected and HIV-uninfected mothers was 3196.6g (95% CI: 3125.3, 3267.9) and 3231.4g (95% CI: 3208.8, 3254.0), respectively. While in 2015-16, the mean birth weight was 3276.4g (95% CI: 3203.6, 3349.1) for children born to HIV-infected mothers and 3221.4g (95% CI: 3199.7, 3243.1) for children born to HIV-uninfected mothers. Perinatal mortality was recorded as 38 deaths per 1000 live births and 35 deaths per 1000 live births in 2010 and 2015-16, respectively. In 2010, 11.1%(95% CI: 10.1, 12.0) mothers were HIV-infected and 88.9%(95% CI: 88.0,90.0) mothers were HIV-uninfected and in 2015-16, 9.2%(95% CI: 8.4,10.1) mothers were HIV-infected and 91.2%(95% CI: 90.4, 91.9) mothers were HIV-uninfected. Perinatal mortality was recorded as 55 per 1000 live births among HIV-infected mothers and 40 per 1000 live births among HIV-uninfected mothers in 2010. In 2015-16, perinatal mortality was recorded as 37 per 1000 live births among HIV-infected mothers and 31 per 1000 live births among HIV uninfected mothers.

Maternal HIV-infection data collection

For both the 2010 and 2015-16 MDHS, interviewers collected finger-prick blood specimens from women age 15-49 years and men age 15-54 years who consented to laboratory HIV testing. The protocol for blood specimen collection and analysis was based on the anonymous linked protocol developed for the DHS Program. This protocol allows for the merger of HIV test results with the socio-demographic data collected in the individual questionnaires after removal of all information that could potentially identify an individual. Interviewers explained the procedure, the confidentiality of the data, and the fact that the test results would not be made available to respondents [36, 37]. A sub-sample of 7924 and 8271 women were tested for HIV using an ELISA test during the 2010 and 2015-16 MDHS, respectively. The data set used in the analysis was children data which was based on woman and household questionnaires.

Statistical analysis

The maternal HIV infection effect was estimated for birth weight and perinatal mortality by comparing the outcomes $ATE = E[Y_i(1) - Y_i(0)]$ between HIV-infected and HIV-uninfected mothers across covariates that were considered as possible confounders using standard statistical balancing tools. The distribution of the covariates must not differ across the two HIV groups, so that the effect on birth weight or on perinatal mortality should be attributed to HIV infection only, if no other unmeasured confounding is present. Three approaches that balance confounders were assessed and the best method was chosen to estimate the effect of maternal HIV infection on birth weight and perinatal death. A brief description of three matching methods; a combination of exact and nearest neighbour matching, the propensity score matching and the inverse probability weighting on the PS are presented in the following subsections. The balancing performance of the three methods is assessed using the standardised mean difference (SMD). The method that yielded the best balance was then used to assess the association between HIV status on birth weight and perinatal mortality.

Identifying confounders from the data

Several variables were identified from the MDHS data as potential confounders for the effect of maternal HIV on birth weight. These include; region, place of residence, wealth, age, maternal anaemia, smoking status of a mother, education level, size of child at birth, anaemia, ever terminated a pregnancy, mothers' body mass index (BMI) and marital status. For the effect of maternal HIV infection on perinatal mortality, we identified residence, place of delivery, maternal anaemia, size of child at birth, ever terminated pregnancy, body mass index (BMI), education of a mother and wealth as potential confounders. A list of the variables, that significantly affected HIV infection status and the pregnancy outcomes for both the 2010 and 2015-16 data, are shown in Table 8 in the Appendix.

Handling missing data

The MDHS data had some missing data on important variables and hence it was essential to control for missing data. We initially did a visualisation of the missing pattern among our variables of interest as displayed in Table 7 in the Appendix. We noted that the data had more than 10% of missing data. Hence, taking into consideration that severity may have been affected by missing data, a multiple imputation on the missing data was done for both the 2010 and 2015-16 MDHS dataset. The mice package in R was used. All further analysis were done on the imputed data.

Exact and nearest neighbour (NN) matching

Exact matching was used to match observed covariates between HIV-infected and HIV-uninfected mothers. If we let i to represent unmatched HIV-infected mothers and j to be a set of unmatched HIV-uninfected mothers, then exact matching ensures that for measured covariates X, $X_i = X_j = x$ [31, 38]. The exact matching was combined with the traditional nearest neighbour matching, which

considers the distance metrics between covariates, [39] to address the problem of increase in dimensionality that comes with exact matching only on the covariates [40]. The 1 : 3 nearest neighbour matching was used to select for each HIV-infected mother *i*, three corresponding HIV-uninfected mothers *j* with the smallest distance between them. HIV-infected and HIV-uninfected mothers were matched on the covariates region, residence, anaemia, age, body mass index (BMI), size of child at birth and wealth using the NN matching for both datasets. Exact match was done on covariates region, ever terminated pregnancy and mothers' age.

Matching based on the propensity score (PS)

An alternative to the traditional matching method between HIV-infected and HIV-uninfected on the covariates is known as a propensity score. The propensity score [27, 41] is the probability that a subject is HIV-infected conditional on observed covariates, denoted as;

$$e(x_i) = P(\text{HIV} = 1 \mid x_i) \tag{1}$$

The propensity score was estimated using a multiple binary logistic regression with exposure (HIV status = 1 or 0) as an outcome variable against potential confounders. The propensity score, known as a balancing score, can produce unbiased average treatment effects [27] under a common assumption known as the strong ignorability assumption. The model is as follows;

$$\begin{aligned} P(\text{HIV Status} = 1 \mid X_1) &= \log\left(\frac{\pi(x)}{1-\pi(x)}\right) \\ &= \beta_0 + \beta_1 \text{Region} + \beta_2 \text{mothers age} \\ &\quad + \beta_3 \text{residence} + \beta_4 \text{Wealth} \\ &\quad + \beta_5 \text{pregnancy termination} \\ &\quad + \beta_6 \text{anaemia} + \beta_7 \text{birth size} + \beta_8 \text{BMI} \end{aligned}$$

The vector of parameters β were estimated based on the whole sample. $\pi(x)$ is the probability of a mother being HIV-infected ($Y = 1$). Subjects were 1 : 1 matched without replacement on the logit of the propensity using Austin's [42] recommended caliper of 0.2, as it is known to produce low mean square error (MSE).

Inverse probability weighting (IPW) based on the PS

The inverse probability weight has been known to perform better in controlling confounding unlike matching on the propensity score [34, 38]. The PS weighting re-weights HIV-infected and HIV-uninfected mothers by creating a population that is free from measured bias [28]. The inverse of the probability of a mother being HIV-infected, conditional on covariates in the data (inverse of the PS) is given by; $w_i = \frac{1}{e(x_i)}$ and the inverse of the probability of a mother being HIV-uninfected, conditional on covariates is given by $w_i = \frac{1}{1-e(x_i)}$, where $e(x_i)$ is the propensity score. To estimate the HIV effect based on the weights, we defined the inverse probability weights with respect to HIV-infection status as;

$$w_i = \frac{A_i}{e(x_i)} + \frac{(1-A_i)}{1-e(x_i)}, A_i = \text{HIV status} = 0, 1 \tag{2}$$

The IPW was estimated as a weight on the PS using the p-scores obtained for mothers' HIV status on confounders, place of residence, education, wealth quintile, mothers' age, employment, whether a mother ever terminated a pregnancy, BMI, region, size of child at birth, BMI, marital status and maternal anaemia. To estimate the HIV effect on birth weight and on perinatal mortality using IPW, a weighted linear regression model and a weighted logistic regression model was fit using the inverse probability weights, respectively.

Balance of measured covariates

Statistical comparison among the balancing methods

The balancing methods were assessed in their ability to balance the distribution of measured potentially confounded covariates between HIV-infected and HIV-uninfected mothers by reducing the standardised mean difference. The method that produced the smallest absolute standardised difference was then used to estimate the association of maternal HIV-infection on the outcomes. The standardised differences compared the prevalence [43] of the categorical covariates in HIV-infected and HIV-uninfected mothers, across the measured covariates. Even though there is no firm agreement on what value of the standardised difference denotes imbalance between exposed and unexposed subjects in the matched sample, the study adopted the decision criterion from some researchers who proposed that a standardised difference of 0.1 (10 per cent) denotes meaningful balance in the measured covariates [44].

In addition to balance checking using the standardised differences, we tested for balance by sub-classifying the propensity scores into quintiles. The full sample (HIV-infected mothers) was separated into five quintiles defined by their propensity scores. For the covariates, frequencies were compared before and after adjusting for propensity score quintile among HIV-infected and HIV-uninfected mothers. Specifically, the p-values for HIV status across the covariates were compared before and after adjustment for propensity score quintile to determine whether balance on the covariates was achieved.

Moderator analysis

Moderation analysis has been used in various public health disciplines including studies on HIV [45, 46]. Most studies have examined the moderating effect of a variable by considering simple interaction effects [45, 47], while

others have done moderation analysis using complex statistical methodologies such as the generalised estimating equations [46].

In a full moderation setting, one needs all information of maternal HIV-infection between different levels of ART status. However, in our study, we are limited in that ART status was available for only those mothers who were HIV-infected. Therefore, we did a partial moderation analysis that considered the association between ART status and birth weight and perinatal mortality among HIV-infected mothers on ART and HIV-infected mothers not on ART. We conducted an analysis of variance (ANOVA) to test the difference in mean birth weight between HIV-uninfected mothers, HIV-infected mothers on ART and HIV-infected mothers not on ART. The Bonferroni ad hoc test was used to measure the difference in means. We further tested the difference in proportion of perinatal mortality between HIV-uninfected mothers, HIV-infected mothers on ART and HIV-infected mothers not on ART. We used a logistic regression to assess the effect of ART on perinatal mortality as a moderating variable.

Assumptions to validate HIV-infection effect

In order to identify the effect of HIV-infection on birth weight and perinatal mortality from cross-sectional data, several assumptions must be considered. The first two unverifiable assumptions are consistency and exchangeability (strongly ignorable assumption). The consistency assumption states that the observed outcome denoted as Y_i is equal to an individuals potential outcome $Y_i(a)$, $Y_i = Y_i(a)$. The exchangeability assumption states that, conditional on measured covariates, the potential outcomes are independent of the exposure $(Y(1), Y(0)) \amalg A|X$. This implies that once we control for the measured covariates, we reduce confounding. Therefore, we need to ensure that all covariates that affect maternal HIV status, birth weight and perinatal mortality are controlled for. The positivity assumption states that each subject must have a non-zero probability of being either HIV-infected or HIV-uninfected. An additional assumption is the Stable Unit Treatment Value Assumption (SUTVA) which assumes independence in the data between the different subjects. Since the design for the MDHS data follows a multi-stage cluster sampling, we assume there was no interdependence between clusters and because the subjects were randomly selected, hence, reducing selection bias.

Sensitivity analysis

A sensitivity analysis was performed to measure the level of unobserved confounders that would affect the obtained estimates. The MDHS data did not include variables that were likely to influence birth weight and perinatal mortality. Variables such as, clinical factors, mothers diet, fetal height, were not available in the MDHS and were considered as unmeasured confounders. A sensitivity analysis method developed by Rosenbaum et al. [48] was used. The Hodges-Lehmann test and Wilcoxon Signed Rank test were used to conduct the sensitivity analysis. The maximum value of Γ was set to 1.3 with 0.05 increments. No unmeasured confounders implies that $\Gamma = 1$. The rbounds package in R was used to obtain the upper and lower bounds on the *p*-value and point estimate [49].

Results

A summary of the distribution of mothers' baseline characteristics by mothers' HIV status are shown in Table 1. Valid data was available for 4,111 and 4,759 under-five children whose mothers had an HIV test result, for the 2010 and 2015 MDHS, respectively. Analysis was done to test the effect of maternal HIV-infection on birth weight and perinatal mortality. From the 2010 data, 316 mothers were HIV-infected, 95 (49.5%) of them were on ART and 248 (79.5%) were residing in rural areas. Forty (12.7%) HIV-infected mothers were from the southern region, 179(56.7%) were aged 25 to 34 years and a third of them (n=103, 32.6%) were from poor households with 196(62.0%) having had a primary education. One hundred and twenty (38.7%) HIV-infected mothers had anaemia, 52(16.5%) reported to have ever terminated a pregnancy and 192(60.8%) were married. In 2015-16, 234(9.0%) of the mothers were HIV-infected of whom 168(71.8%) were from the rural area and 105(44.9%) were aged between 25 to 34 years old. Thirty nine (16.7%) of the HIV-infected mothers were in the middle class category, 144(61.5%) had a primary education and 167(71.4%) were married. One hundred and eighteen (50.4%) of the HIV-infected mothers had anaemia and 38(16.2%) reported to have ever terminated a pregnancy.

Comparison before and after balancing confounders across HIV groups

Before matching, for both datasets, mothers' demographic and health characteristics, region, place of residence, marital status, age of a mother, maternal anaemia, marital status and whether a mother ever terminated a pregnancy varied among HIV-infected and HIV-uninfected mothers as presented in Table 1. Wealth quintile was different among HIV-infected and HIV-uninfected mothers for the 2010 data only.

After applying the balancing methods on the covariates for the 2010 and 2015-16 datasets, no significant difference was observed for the covariates between the HIV-infected and HIV-uninfected, on the matched sample, as shown in Table 9 in the Appendix. Before balancing the differences, the absolute value of standardised difference was greater than 0.1 across the covariates. After balancing the covariates using the three methods, the absolute standardised difference for the three methods

Table 1 Distribution of confounders between HIV-infected and HIV-uninfected mothers before matching for MDHS 2010 and 2015-16

Characteristic	HIV status (2010)			HIV Status(2015)		
	HIV-infected(%)	HIV-uninfected(%)	*p*-value	HIV-infected(%)	HIV-uninfected(%)	*p*-value
Overall	316(11.0)	2681(89.0)		234(9.0)	2456(91.0)	
Residence						
Urban	68(21.5)	320(11.9)		66(28.2)	409(16.7)	
Rural	248(78.5)	2361(88.1)	<0.001	168(71.8)	2047(83.4)	<0.001
Region						
Northern	40(12.7)	581(21.7)		30(12.8)	481(19.6)	
Central	76(24.1)	933(34.8)		49(20.9)	887(36.1)	
Southern	200(63.3)	1167(43.5)	<0.001	155(66.2)	1088(44.3)	<0.001
Age						
15-24	72(22.8)	983(36.7)		62(26.5)	1086(44.2)	
25-34	179(56.7)	1198(44.7)		105(44.9)	1025(41.7)	
35 above	65(20.6)	500(18.7)	<0.001	67(28.6)	345(14.1)	<0.001
Wealth						
Poor	103(32.6)	1004(37.5)		90(38.5)	1054(42.9)	
Middle	60(19.0)	609(22.7)		39(16.7)	478(19.5)	
Rich	153(48.4)	1068(39.8)	0.013	105(44.9)	924(37.6)	0.09
Education						
None	46(14.6)	322(12.0)		25(10.7)	240(9.8)	
Primary	196(62.0)	1829(68.2)		144(61.5)	1582(64.4)	
Secondary	70(22.2)	508(19.0)		62(26.5)	585(23.8)	
Tertiary	4(1.2)	22(0.8)	0.154	3(1.3)	49(2.0)	0.647
Anaemia						
No	190(61.3)	1939(74.2)		116(49.6)	1757(71.6)	
Yes	120(38.7)	675(25.8)	<0.001	118(50.4)	690(28.5)	<0.001
Ever terminated pregnancy						
No	264(83.5)	2376(88.6)		196(83.8)	2235(91)	
Yes	52(16.5)	304(11.3)	0.028	38(16.2)	221(9.0)	<0.001
Marital Status						
Never married	10(3.2)	78(2.9)		5(2.0)	119(4.9)	
Married	192(60.8)	2050(76.5)		167(71.4)	1928(78.5)	
Living together	30(9.5)	282(10.5)		13(5.6)	164(6.7)	
Widowed	20(6.3)	39(1.5)		10(4.3)	24(1.0)	
Divorced	39(12.3)	117(4.4)		18(7.7)	93(3.8)	
Not Living together	25(7.9)	115(4.3)	<0.001	21(9.0)	128(5.2)	<0.001
Place of delivery						
Home	13(4.4)	102(4.1)		5(2.3)	32(1.4)	
Government Hosp.	247(83.5)	2047(82.7)		200(90.9)	2051(88.4)	
Private	8(2.7)	71(2.9)		4(1.8)	58(2.5)	
Christian/Mission Hosp.	28(9.5)	255(10.3)	0.96	11(5.0)	180(7.8)	0.301
Last birth C/section						
No	294(93.0)	2520(94)		215(92.3)	3263(92.6)	
Yes	21(6.7)	158(5.9)	0.55	18(7.7)	180(7.4)	0.675

Table 2 Absolute standardised differences for before and after balancing the covariates with the three methods for the 2010 MDHS

Covariates	Unmatched	NN Match	PS Match	IPW
Residence	0.29	0.00	0.04	0.06
Wealth	0.13	0.05	0.03	0.00
Age	0.16	0.04	0.06	0.01
Anaemia	0.28	0.05	0.03	0.02
Region	0.37	0.11	0.03	0.03
Education	0.05	0.01	0.05	0.02
Marital status	0.02	0.03	0.02	0.00
Size at birth	0.05	0.02	0.03	0.04
Ever terminated pregnancy	0.15	0.01	0.1	0.04

was less than 0.1 as shown in Tables 2 and 3. We further used the standardised differences obtained for the three methods to choose the balancing method that produced the best balance among the covariates and used it to estimate the effect of maternal HIV infection on birth weight and perinatal mortality. For both the 2010 and 2015-16 data, the IPTW gave smaller absolute standardised differences as compared to the PS match and the NN match method.

To further ensure that balance was obtained, we stratified the matched observations into quintiles according to their propensity scores. We observed that, within each quintile, there was no significant difference between HIV status for most of the covariates identified as confounders. These results are shown in Tables 10 and 11 in the Appendix. This demonstrated that balance was achieved.

Maternal HIV effect estimation

Table 4 presents results for the effect of maternal HIV-infection on birth weight and perinatal mortality using the inverse probability weights. The results for the 2010 data showed that maternal HIV-infection was negatively associated with birth weight. Infants of HIV-infected mothers had a lower average birth weight of -25.3g (95% CI: -95.5, -7.4) as compared to infants of HIV-uninfected mothers. The 2015-16 results showed a significant positive association between maternal HIV infection status and birth weight. Infants of HIV-infected mothers had higher average birth weight 116.3g (95% CI: 27.8, 204.7) compared to infants of HIV-uninfected mothers. In addition, a comparison of the odds ratio of perinatal mortality between HIV-infected and HIV-uninfected mothers showed an increase in perinatal deaths in HIV-infected mothers in 2010. The odds of perinatal mortality was 1.5 times greater among infants of HIV-infected mothers compared to infants of HIV-uninfected mothers. However, in 2015, there was no significant difference in perinatal mortality between HIV-infected and HIV-uninfected mothers.

Table 5 presents the effect of ART status on the association between maternal HIV-infection on birth weight and perinatal mortality among mothers who were HIV-infected. There was no difference in birth weight between infants born from HIV-infected mothers on ART and those born to HIV-infected mothers not on ART. However, a significant difference (p-value= 0.02) in birth weight was observed between infants born to HIV-uninfected mothers and those born to HIV-infected mothers on ART as shown in Table 5. Infants of HIV-uninfected mothers had a higher average birth weight of 3231.3g (95% CI: 3208.8, 3253.9) as compared to the

Table 3 Absolute standardised differences for before and after balancing the covariates with the three methods for the 2015-16 MDHS

Covariates	Unmatched	NN Match	PS Match	IPW
Residence	0.30	0.01	0.02	0.01
Wealth	0.14	0.03	0.01	0.06
Age	0.39	0.10	0.03	0.06
Anaemia	0.43	0.02	0.06	0.01
Region	0.35	0.19	0.06	0.03
Education	0.02	0.01	0.06	0.01
Marital status	0.28	0.03	0.01	0.01
Size at birth	0.10	0.02	0.04	0.02
Ever terminated pregnancy	0.16	0.02	0.09	0.03

Table 4 The effect of maternal HIV-infection on birth weight and perinatal mortality

	Child Health Outcomes 2010		**Child Health Outcomes 2015-16**	
	Birth Weight	Perinatal Mortality	Birth Weight	Perinatal Mortality
	Mean.Diff(95% CI)	OR(95% CI)	Mean.Diff(95% CI)	OR(95% CI)
Unadjusted	-97.2(-160.4, -34)	1.8(1.2, 2.8)	52.6 (41.6, 125.5)	1.4(0.8, 2.6)
Adjusted weights	-25.3(-95.5, -7.4)	1.5(1.1, 3.1)	116.3(27.8, 204.7)	1.0(0.4, 1.6)

Variables age, residence, region, wealth, marital status, ever terminated a pregnancy, anaemia and body mass index were included in the propensity score model

average birth weight of 3128.1g (95%: 2982.5, 3273.7) for infants born to HIV-infected mothers on ART.

After observing a significant difference in prevalence of perinatal mortality between HIV-infected mothers on ART and those not on ART, we assessed the moderating effect of ART on the association between maternal HIV infection and perinatal mortality. There was a significant difference in perinatal mortality between HIV-infected mothers on ART and HIV-infected mothers not on ART as presented in Table 6. The odds of perinatal mortality was 3.9 times greater among infants of HIV-infected mothers not on ART as compared to infants of HIV-infected mothers on ART. In addition, there was no difference in perinatal mortality between HIV-infected mothers on ART and HIV-uninfected mothers.

Sensitivity analysis

A sensitivity analysis was done after matching to observe if the estimates obtained were sensitive to bias from unmeasured confounders. Table 12 in the Appendix presents results on the upper and lower bounds on the Wilcoxon Signed Rank *p*-value and the upper and lower bounds on the Hodges-Lehmann point estimate for birth weight. Based on the tables, the results for the sensitivity analysis showed that the *p*-value for the estimate of birth weight between HIV-infected and HIV-uninfected mothers with similarly measured covariates, would change with a small bias influenced by unmeasured confounders. The Hodges-Lehmann estimates similarly showed that a small bias influenced by unobserved differences would cause a change in the estimates.

Discussion

The study set out to assess the effect of maternal HIV infection on birth weight and perinatal mortality and the effect of ART on these pregnancy outcomes. The study has shown that, in 2010, maternal HIV infection had an adverse effect on birth weight and perinatal mortality. However, in 2015-16, infants of HIV-infected mothers significantly weighed more than infants born to HIV-uninfected mothers. ART status had no effect on the association between maternal HIV infection and birth weight. However, ART did have a positive effect on the association between maternal HIV infection and perinatal mortality.

These findings from our study are similar to previous studies that reported a high risk of low birth weight among HIV-infected women [4, 9, 15, 50]. This has been attributed to the fact that HIV-infected women not on treatment have suppressed immune system which may have affected child growth and contributed to LBW among their infants [1, 4, 18]. Consistent with previous results [8, 18, 51], the study has also shown that perinatal mortality was higher among infants of HIV-infected mothers compared to those of HIV-uninfected mothers. These findings are in line with studies that have reported that LBW and perinatal mortality are interrelated and that LBW accounts for the majority of perinatal deaths [18, 51]. Since low birth weight was high among infants of HIV-infected mothers, we postulate that most of them did not survive after the first week from birth.

There has been a lack of consensus on the effect of maternal HIV infection on birth weight with some studies attributing LBW to maternal HIV infection [4, 15, 17], while others attributing it to the use of ART [2, 8, 14, 16]. The findings from the current study, which suggest that the association between maternal HIV infection and birth weight was not affected by ART uptake, is consistent with findings of Xiao et al. [4] who found that ART exposure

Table 5 Mean Birth weight and proportion of perinatal deaths for HIV-uninfected and by ART status among HIV-infected mothers for the 2010 MDHS data

	Birth weight		**Perinatal Mortality**	
	Mean (95% CI)	***p*-value**	**Prop.(95% CI)**	***p*-value**
HIV-infected on ART (m_1)	3128.1(2982.5,3273.7)	0.64 (m_1 vs m_2)	4.3(0.2, 8.3)	0.03(m_1 vs m_2)
HIV-infected not on ART(m_2)	3031.9(2865.8, 3198.0)		9.4(3.5, 15.2)	
HIV-uninfected (m_3)	3231.4(3208.8, 3253.9)	0.02(m_1 vs m_3)	4.0(3.4, 4.7)	0.9(m_1 vs m_3)

Table 6 The effect of ART status on perinatal mortality

	Perinatal Mortality	
	OR	**95% CI**
HIV-infected on ART	**	
HIV-infected not on ART	3.9	(1.1, 14.8)
HIV-uninfected	0.7	(0.2, 2.5)

** - reference category
Variables age, residence, region, wealth, marital status, ever terminated a pregnancy, anaemia and body mass index were included in the propensity score model

did not decrease or increase the risk of LBW in HIV-infected women. However, Gibango et al. [18] reported that HIV-positive mothers not on ART were likely to deliver an infant with low birth weight. Even though ART did not mitigate the association between maternal HIV infection and birth weight, it did mitigate the association between maternal HIV and perinatal mortality. Contrary to our findings, previous studies have shown that adverse pregnancy outcomes such as stillbirths are high among HIV-infected women on ART [52]. Lower perinatal deaths among HIV-infected mothers on ART may not entirely be as a result of the ART uptake but may also include uptake of good nutritional diet by the mothers. PMTCT programs through antenatal care (ANC) services ensure HIV-infected pregnant mothers are on ART and encourage good maternal nutrition and infant feeding practices. However, the 2010 MDHS report revealed that only 12% and 48% of pregnant women had their first ANC visit during the first trimester and during the fourth and fifth month of pregnancy, respectively [36]. PMTCT interventions through ANC clinics and community ART groups [53] may have been beneficial in improving perinatal mortality among infants of HIV-infected women but did not help in improving birth weight in 2010. Perhaps the benefits of the PMTCT programmes provided at ANC clinics during this period could have been more pronounced if frequency and timing of ANC visit was high.

In 2015-16, there was an increase in birth weight among infants born to HIV-infected mothers compared to infants born to HIV-uninfected mothers. This could be attributed to the fact that in July 2011, Malawi was one of the first countries to implement Option B+ approach as an effort to deal with HIV/AIDS and its impact. This is an intervention where all pregnant women living with HIV were offered ART for life irrespective of their CD4 count [54]. Reports show that with the introduction of option B+, there was a huge improvement in PMTCT in Malawi [55]. The estimated percentage of pregnant women living with HIV who received ART drugs for PMTCT in Malawi was 84% in 2016 [56]. Perhaps the HIV-infected women were part of this intervention and it may have helped them boost their immunity [16, 55] which in turn resulted in good child growth. The findings of the current study further showed that perinatal mortality was similar between HIV-infected and HIV-uninfected. The results may suggest that since infants of HIV-infected mothers had higher birth weight then they were less likely to die of LBW or prematurity. In addition, through antenatal care services, HIV-infected pregnant mothers are put on ART and are encouraged to attend at least four ANC visits where they receive the required iron supplements, folic acid and Ready to Use Therapeutic Food (RUTF) [57]. This may have helped HIV-infected pregnant mothers to maintain a balanced nutritional status which in turn helped in child growth as compared to HIV-uninfected mothers who only receive iron supplements and folic acids without additional food supplements. Previous studies have also shown that prenatal and antenatal care visits among pregnant mothers result in a decrease in perinatal mortality [58, 59]. The findings from our study may suggest a major improvement in the implementation of PMTCT interventions by the Ministry of Health and other key stakeholders.

Table 7 Missing Pattern by studied sample

	2010 MDHS data		2015 MDHS data	
Variable	**Missing**	**Percent**	**Missing**	**Percent**
Birth weight	806	19.1	484	10.2
Terminated pregnancy	423	10.0	508	10.7
Residence	414	10.0	455	9.5
HIV-infection status	317	9.0	480	10.1
Size at birth	408	9.9	473	10.0
Anaemia	522	13.0	448	9.4
age	415	10.0	480	10.1
Marital status	415	10.0	463	9.7
perinatal deaths	434	10.6	739	15.0

Strengths and limitations

One of the strengths of the study is that data originating from a large survey was used and hence, using a rich data with a vast number of socio-demographic characteristics of participants. The study also used matching and weighting methods to reduce confounding for the observed covariates. Several limitations of the study must be addressed. The study used survey data that collected retrospective information on births and deaths. Underreporting of births and deaths for children who were not living at the time of the survey may be common. Mothers may have been reluctant to talk about their dead children either because the subject brought back sad memories or because their culture discouraged mentioning of the dead. Coupled to that, the analysis was done on cross-sectional data and despite controlling for confounding, it was difficult to assess the causal effect of maternal HIV status on birth weight and on perinatal and infant mortality, as the data is very likely to have unmeasured covariates that may act as confounders. In addition, it is difficult to measure the change in birth weight among HIV-infected mothers over a certain period of time. Furthermore, some important confounding risk factors that are known to affect birth weight were not available in the data. These include; unfavourable genetic factors which predispose infants to be smaller in size and length [26], maternal nutrition, maternal passive smoking, drinking, folic acid supplement use, taking drugs and family history of birth defects [50]. Therefore, the obtained estimates may be sensitive to unmeasured confounders. In addition, knowledge of ART timing, initiation of ART and type of ART regimen were not available in the data. These are known to have influenced the risk of adverse birth outcomes in HIV-infected mothers. On the other hand, knowledge on date of HIV acquisition was not available in the data hence it may be possible that a mother was infected with HIV

Table 8 Confounder identification

Variable	**Outcome**	**Exposure**	
	Birth weight (*p*-value)	Perinatal death(*p*-value)	HIV status (*p*-value)
Age of a mother	<0.001	<0.001	<0.001
Region	NA	0.006	<0.001
Residence	0.016	0.946	<0.001
Education	0.011	0.43	0.012
Size of child at birth	<0.001	0.06	0.31
Ever had terminated pregnancy	0.845	<0.001	0.007
Anaemia	<0.001	0.90	<0.001
Marital status	NA	NA	<0.001
place of delivery	NA	0.249	0.97
Wealth	0.97	0.286	0.008

(a) Confounders that influence HIV and the outcomes for the 2010 MDHS

	Birth weight (*p*-value)	Perinatal death(*p*-value)	HIV status (*p*-value)
Age of a mother	<0.001	0.226	<0.001
Region	0.133	0.35	<0.001
Residence	0.49	0.126	<0.001
Education	0.096	0.716	<0.001
BMI	0.017	NA	0.577
Size of child at birth	<0.001	<0.001	0.304
Ever had terminated pregnancy	0.730	<0.001	0.001
Anaemia	0.21	0.50	<0.001
Marital status	NA	NA	<0.001
Place of delivery	NA	0.26	0.41
Wealth	0.001	0.287	0.017

(b) Confounders that influence HIV and the outcomes for the 2015-16 MDHS

Confounder Identification A verification of the variables was done by fitting a linear regression model for the effect of each confounder on child birth weight and did a crosstabulation using Pearsons Chi-square test for the effect of the confounders on perinatal mortality. We also checked the association between the identified variables and mothers HIV-infection status. Variables that were not known to influence either HIV status or birth weight or perinatal death were labelled as "not applicable"

after infant birth. In this case, HIV status would not affect birth weight. Furthermore, ART status was not available within the 2015-16 MDHS data. Hence, conclusions about the effect of ART status on birth weight can only be attributed to the 2010 data and may not explain changes in birth weight for the 2015-16 data. The assumptions stated to estimate the maternal HIV effect on the outcomes were not tested, hence estimates may be prone to bias.

Conclusion

Our findings suggest that maternal HIV infection had an impact on birth weight and perinatal mortality. The uptake of ART had a moderating effect on the association between maternal HIV-infection and perinatal mortality with higher rates among HIV-infected mothers not on ART. The study has also shown a significant increase in birth weight between the 2010 and 2015-16 MDHS. In addition, the findings show a change in effect of maternal HIV-infection on perinatal deaths between the period 2010 and 2015. These results may be explained by the recent success of policies and interventions under the PMTCT programme such as wide coverage of antenatal care services and the option B+ strategy. Therefore, there is a need to emphasise the

Table 9 Distribution of Confounders between HIV-infected and HIV-uninfected after matching for MDHS 2010 and 2015-16 data

Characteristic	Mothers HIV status (2010)			Mothers HIV Status(2015)		
	HIV-infected	HIV-uninfected	*p*-value	HIV-infected(%)	HIV-uninfected(%)	*p*-value
Overall	301(51.2)	306(48.8)		387(50.1)	386(49.9)	
Residence						
Urban	66(20.6)	58(19.0)		112(28.9)	120(31.1)	
Rural	255(79.4)	248(81)	0.61	275(71.1)	266(68.9)	0.52
Region						
Northern	42(13.1)	37(12.1)		46(11.9)	41(10.6)	
Central	84(26.2)	101(33.0)		91(23.5)	117(30.3)	
Southern	195(60.8)	168(54.9)	0.17	250(64.6)	228(59.1)	0.10
Age						
15-24	78(24.3)	97(31.7)		93(24.0)	90(23.3)	
25-34	173(53.9)	122(39.9)		181(46.8)	197(51.0)	
35 above	70(21.8)	87(28.4)	0.13	113(29.2)	99(25.7)	0.44
Wealth						
Poor	100(31.2)	96(31.4)		144(37.2)	137(35.5)	
Middle	70(21.8)	65(21.2)		65(16.8)	67(17.4)	
Rich	151(47.0)	145(47.4)	0.98	178(46)	182(47.1)	0.88
Education						
None	50(15.6)	41(13.4)		44(11.4)	54(14.0)	
Primary	199(62)	206(67.3)		242(62.5)	213(55.2)	
Secondary	69(21.5)	57(18.6)		94(24.3)	107(27.7)	
Tertiary	3(0.9)	2(0.6)	0.57	7(1.8)	12(3.1)	0.17
Anaemia						
No	193(60.1)	196(64.1)		196(50.7)	193(50.0)	
Yes	128(39.9)	116(35.9)	0.31	191(49.3)	193(50.0)	0.86
Ever terminated pregnancy						
No	272(84.7)	259(84.6)		332(85.8)	322(83.4)	
Yes	49(15.3)	47(15.4)	0.97	55(14.2)	64(16.6)	0.30
Marital Status						
Never married	8(2.5)	3(1.0)		10(2.6)	8(2.1)	
Married	197(61.4)	202(66.0)		273(70.5)	275(71.2)	
Living together	32(10.0)	36(11.8)		18(4.7)	26(6.7)	
Widowed	21(6.5)	5(1.6)		13(3.4)	12(3.1)	
Divorced	37(11.5)	22(7.2)		37(9.6)	23(6.0)	
Not Living together	26(8.1)	38(12.4)	0.03	36(9.3)	42(10.9)	0.36

Table 10 Comparison of Covariates for HIV status Before and After Propensity Score Stratification

Covariate	*p*-value(2010))		*p*-value(2015-16)	
	Before stratification	**After stratification**	**Before stratification**	**After Stratification**
Age				
15-24	<0.001	0.800	<0.001	0.767
25-34	0.009	0.922	0.567	0.316
35 above	0.02	0.893	<0.001	0.14
Wealth				
Poor	0.553	0.440	0.04	0.78
Medium	0.02	0.817	0.34	0.70
Rich	0.01	0.618	0.004	0.56
Residence				
Urban	<0.001	0.424	<0.001	0.25
Rural	<0.001	0.424	<0.001	0.25
Region				
Northern	0.001	0.514	<0.001	0.08
Central	0.001	0.681	<0.001	0.649
Southern	<0.001	0.368	<0.001	0.5
Anemia				
No	0.04	0.537	<0.001	0.21
Yes	0.04	0.537	<0.001	0.21

importance of uptake of PMTCT guidelines for effective child health outcomes. Statistical methods that control for confounders are useful tools to address the lack of randomised control trials in assessing the effect of maternal HIV on pregnancy outcomes based on observational data. We recommend future studies to consider using causal inference techniques when an investigation of association of risk factor and outcome is undertaken using observational studies such as health surveys.

Appendix

Table 11 Quintiles Based on estimated Propensity Scores

Distribution	Sample (N)		*P*-value				
	HIV-uninfected	**HIV-infected**	**age**	**Residence**	**Region**	**Anaemia**	**Ever Terminated pregnancy**
(a) HIV status Group Differences on Covariates age, residence, region, anemia and ever terminated pregnancy, by Strata Based on Propensity Score Quintiles on the matched sample for 2010 MDHS							
Quintile 1	719	23	0.06	0.069	0.688	0.41	0.02
Quintile 2	692	51	0.502	0.653	0.513	0.1	0.1
Quintile 3	683	58	0.347	0.211	0.06	0.95	0.59
Quintile 4	656	86	0.9	0.1	0.39	0.1	0.42
Quintile 5	590	152	0.5	0.97	0.3	0.72	0.22
(b) HIV status Group Differences on Covariates age, residence, region, anemia and ever terminated pregnancy, by Strata Based on Propensity Score Quintiles on the matched sample for 2015-16 MDHS							
Quintile 1	870	24	0.236	0.04	0.67	0.34	0.83
Quintile 2	855	39	0.38	1.00	0.323	0.16	0.37
Quintile 3	821	72	0.06	0.29	0.07	0.07	0.87
Quintile 4	796	98	0.12	0.37	0.04	0.16	0.39
Quintile 5	713	180	0.04	0.903	0.07	0.526	0.19

Table 12 Rosenbaum's Sensitivity Analysis

	***P*-value**		**Estimate**	
Gamma	**Lower bound**	**Upper bound**	**Lower**	**Upper**
(a) Wilcoxon Signed Rank P-Value and Hodges-Lehmann Point Estimate for 2010 MDHS				
1	0.001	0.00	400	400
1.1	0.0002	0.00	350	450
1.2	0.0001	0.01	300	500
1.3	<0.001	0.01	300	500
1.4	<0.001	0.02	250	550
1.5	<0.001	0.03	250	550
(b) Wilcoxon Signed Rank P-Value and Hodges-Lehmann Point Estimate for 2015-16 MDHS				
1	0.0004	0.00	-83.3	-83.3
1.1	0.0001	0.00	-133.3	-133.3
1.2	<0.001	0.01	-150.1	150.1
1.3	<0.001	0.01	-166.7	166.7
1.4	<0.001	0.03	-183.3	183.3
1.5	<0.001	0.06	-200	200

Abbreviations

ANC: Antenatal clinic; ANOVA: Analysis of variance; ART: Anti-retroviral therapy; HIV: Human immune-deficiency virus; IPW: Inverse probability weights; LBW: Low birth weight; LMIC: Low-middle income countries; Malawi Demographic and health survey; MTCT: Mother to child transmission; NN: Nearest neighbour; PMTCT: Prevention of mother to child transmission; PS: Propensity score; Randomised control trial; SUTVA: Stable unit treatment value assumption; WHO: World health organisation

Acknowledgements

We acknowledge the DHS team for making the MDHS data available for use and we also acknowledge the support we received from SAMRC through the provision of research facilities.

Authors' contributions

HT performed data management, statistical analysis, and wrote the initial draft of manuscript. SM and DS jointly conceived the ideas, reviewed the statistical analysis and helped with the revision of the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported through the DELTAS Africa Initiative, Sub-Saharan Africa Consortium for Advanced Biostatistics Training (SSACABT). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [SSACABT] and the UK government. Special thanks to South Africa Medical Research Council for hosting and supporting my research visits. The views expressed in this publication are those of the author(s) and not those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

Availability of data and materials

The data that support the findings of this study are available upon request from the Demographic and Health Survey (DHS) website. Upon approval, full access is granted to all unrestricted survey datasets.

Ethics approval and consent to participate

The Malawi Health Research Committee determined that ethical approval was not deemed necessary in this study considering the fact that the study used data from a research study already approved by an ethical research committee. According to the 2010 and 2015-16 MDHS report [36, 37], the MDHS study was ethically approved by Malawi Health Research Committee, Institutional Review Board of ICF Macro, Centre for Disease and Control (CDC) in Atlanta, GA, USA and Prevention IRB. Informed consent was obtained from all eligible persons.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Mathematical Sciences, University of Malawi, Zomba, Malawi. [2]Biostatistics Research Unit, South Africa Medical Research Council, Pretoria, South Africa. [3]School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal, Pietermaritzburg, South Africa. [4]Department of Statistics, University of Pretoria, Pretoria, South Africa. [5]Department of Statistics, University of Pennyslvania, Pennyslvania, USA.

Received: 10 January 2020 Accepted: 3 June 2020
Published online: 22 June 2020

References

1. Dreyfuss ML, Msamanga GI, Spiegelman D, Hunter DJ, Urassa EJ, Hertzmark E, Fawzi WW. Determinants of low birth weight among HIV-infected pregnant women in Tanzania. Am J Clin Nutr. 2001;74(6): 814–26.
2. Rollins NC, Coovadia HM, Bland RM, Coutsoudis A, Bennish ML, Patel D, Newell M-L. Pregnancy outcomes in hiv-infected and uninfected women in rural and urban south africa. JAIDS J Acquir Immune Defic Syndr. 2007;44(3):321–8.
3. Claeson M, Gillespie D, Mshinda H, Troedsson H, VC. Bellagio study group on child survival: Knowledge into action for child survival. Lancet. 2003;362:323–7.
4. Xiao P-L, Zhou Y-B, Chen Y, Yang M-X, Song X-X, Shi Y, Jiang Q-W. Association between maternal HIV infection and low birth weight and prematurity: a meta-analysis of cohort studies. BMC Pregnancy Childbirth. 2015;15:246.
5. Ellis J, Williams H, Graves W, Lindsay MK. Human immunodeficiency virus infection is a risk factor for adverse perinatal outcome. Am J Obstet Gynecol. 2002;186(5):903–6.

6. Brocklehurst P, French R. The association between maternal hiv infection and perinatal outcome: a systematic review of the literature and meta-analysis. BJOG: Int J Obstet Gynecol. 1998;105(8):836–48.
7. Ticconi C, Mapfumo M, Dorrucci M, Naha N, Tarira E, Pietropolli A, Rezza G. Effect of maternal hiv and malaria infection on pregnancy and perinatal outcome in Zimbabwe. J Acquir Immune Defic Syndr 1999. 2003;34(3):289–94.
8. Bayou G, Berhan Y. Perinatal mortality and associated risk factors: a case control study. Ethiop J Health Sci. 2012;22(3):.
9. Nazli A, Chan O, Dobson-Belaire WN, Ouellet M, Tremblay MJ, Gray-Owen SD, Arsenault AL, Kaushic C. Exposure to hiv-1 directly impairs mucosal epithelial barrier integrity allowing microbial translocation. PLoS Pathog. 2010;6(4):1000852.
10. Nkhoma ET, Kalilani-Phiri L, Mwapasa V, Rogerson SJ, Meshnick SR. Effect of hiv infection and plasmodium falciparum parasitemia on pregnancy outcomes in Malawi. Am J Trop Med Hyg. 2012;87(1):29–34.
11. World Health Organization, et al. UNICEF-WHO low birthweight estimates: levels and trends 2000–2015. No. WHO/NMH/NHD/19.21. United Nations Children's Fund (UNICEF). 2019.
12. Oza S., Lawn J. E., Hogan D. R., Mathers C., Cousens S. N. Neonatal cause-of-death estimates for the early and late neonatal periods for 194 countries: 2000–2013. Bull World Health Organ. 2014;93:19–28.
13. Joint United Nations Programme on HIV/AIDS. Progress report on the global plan towards the elimination of new HIV infections among children by 2015 and keeping their mothers alive. Geneva: Joint United Nations Programme on HIV/AIDS; 2013.
14. Cotter AM, Garcia AG, Duthely ML, Luke B, O'Sullivan MJ. Is antiretroviral therapy during pregnancy associated with an increased risk of preterm delivery, low birth weight, or stillbirth?. J Infect Dis. 2006;193(9):1195–201.
15. Szyld EG, Warley EM, Freimanis L, Gonin R, Cahn PE, Calvet GA, Duarte G, Melo VH, Read JS, Group NPS, et al. Maternal antiretroviral drugs during pregnancy and infant low birth weight and preterm birth. Aids. 2006;20(18):2345–53.
16. Marazzi MC, Palombi L, Nielsen-Saines K, Haswell J, Zimba I, Magid NA, Buonomo E, Scarcella P, Ceffa S, Paturzo G, et al. Extended antenatal use of triple antiretroviral therapy for prevention of mother-to-child transmission of hiv-1 correlates with favorable pregnancy outcomes. Aids. 2011;25(13):1611–8.
17. Tuomala RE, Watts DH, Li D, Vajaranant M, Pitt J, Hammill H, Landesman S, Zorrilla C, Thompson B, et al. Improved obstetric outcomes and few maternal toxicities are associated with antiretroviral therapy, including highly active antiretroviral therapy during pregnancy. JAIDS J Acquir Immune Defic Syndr. 2005;38(4):449–73.
18. Gibango N, Mda S, Ntuli T. Factors associated with delivering premature and/or low birth weight infants among pregnant hiv-positive women on antiretroviral treatment at dr george mukhari hospital, south africa. South Afr J Infect Dis. 2018;33(2):42–5.
19. Chen JY, Ribaudo HJ, Souda S, Parekh N, Ogwu A, Lockman S, Powis K, Dryden-Peterson S, Creek T, Jimbo W, et al. Highly active antiretroviral therapy and adverse birth outcomes among hiv-infected women in botswana. J Infect Dis. 2012;206(11):1695–705.
20. Bian Y, Zhang Z, Liu Q, Wu D, Wang S. Maternal risk factors for low birth weight for term births in a developed region in china: a hospital-based study of 55,633 pregnancies. J Biomed Res. 2013;27(1):14.
21. Khan M, Arbab M, Murad M, Khan M, Abdullah S. Study of factors affecting and causing low birth weight. J Sci Res. 2014;6(2):387–94.
22. Nascimento LFC, Moreira DA. Are environmental pollutants risk factors for low birth weight?. Cad Saúde Pública. 2009;25:1791–6.
23. Gebremedhin M, Ambaw F, Admassu E, Berhane H. Maternal associated factors of low birth weight: a hospital based cross-sectional mixed study in tigray, northern ethiopia. BMC Pregnancy Childbirth. 2015;15(1):222.
24. Ngwira A, Stanley CC. Determinants of low birth weight in malawi: Bayesian geo-additive modelling. PloS One. 2015;10(6):0130057.
25. Metgud CS, Naik VA, Mallapur MD. Factors affecting birth weight of a newborn–a community based study in rural karnataka, india. PloS One. 2012;7(7):40040.
26. Dubois L, Kyvik KO, Girard M, Tatone-Tokuda F, Pérusse D, Hjelmborg J, Skytthe A, Rasmussen F, Wright MJ, Lichtenstein P, et al. Genetic and environmental contributions to weight, height, and bmi from birth to 19 years of age: an international study of over 12,000 twin pairs. PLOS One. 2012;7(2):30153.
27. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41–55.
28. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med. 2004;23(19):2937–60.
29. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656–64.
30. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effect of Confounding in Observational. Multivar Behav Res. 2011;46(3): 399–424. https://doi.org/10.1080/00273171.2011.568786.
31. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat. 1985;39(1):33–8.
32. He H, Hu J, He J. Overview of propensity score methods. In: Statistical Causal Inferences and Their Applications in Public Health Research. Springer International Publishing; 2016. p. 29–48.
33. Sainani KL. Propensity scores: uses and limitations. PM&R. 2012;4(9):693–7.
34. Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-ami statin use. Stat Med. 2006;25(12):2084–106.
35. d'Agostino RB. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med. 1998;17(19):2265–81.
36. National Statistical Office (NSO) and ICF Macro. Malawi Demographic and Health Survey 2010. Zomba, Malawi, and Calverton, Maryland, USA: NSO and ICF Macro; 2011.
37. National Statistical Office (NSO) [Malawi] and ICF. Malawi Demographic and Health Survey 2015–16. Zomba, Malawi, and Rockville, Maryland, USA: NSO and ICF; 2017.
38. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivar Behav Res. 2011;46(3):399–424.
39. Rubin DB. Matching to remove bias in observational studies. Biometrics. 1973;29:159–83.
40. Stuart EA. Matching methods for causal inference: A review and a look forward. Stat Sci Rev J Inst Math Stat. 2010;25(1):1.
41. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. J Am Stat Assoc. 1984;79(387): 516–24.
42. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat. 2011;10(2):150–61.
43. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083–107.
44. Normand S-LT, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, McNeil BJ. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. J Clin Epidemiol. 2001;54(4):387–98.
45. Kim SC, Messing S, Shah K, Luque AE. Effect of highly active antiretroviral therapy (haart) and menopause on risk of progression of cervical dysplasia in human immune-deficiency virus-(hiv-) infected women. Infect Dis Obstet Gynecol. 2013;2013:1–8. https://doi.org/10.1155/2013/784718.
46. Fabian KE, Huh D, Kemp CG, Nevin PE, Simoni JM, Andrasik M, Turan JM, Cohn SE, Mugavero MJ, Rao D. Moderating factors in an anti-stigma intervention for african american women with hiv in the united states: A secondary analysis of the unity trial. AIDS Behav. 20191–11.
47. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51(6):1173.
48. Rosenbaum PR. Observational studies. In: Observational Studies. Springer-Verlag New York; 2002. p. 1–17.
49. Rosenbaum PR. Sensitivity Analysis in Observational Studies. In: Everitt BS, Howell DC, editors. Encyclopedia of Statistics in Behavioral Science, vol 4; 2005. p. 1809–14.
50. Guo L, Qu P, Zhang R, Zhao D, Wang H, Liu R, Mi B, Yan H, Dang S. Propensity score-matched analysis on the association between pregnancy infections and adverse birth outcomes in rural northwestern china. Sci Rep. 2018;8(1):5154.

51. Naniche D, Bardají A, Lahuerta M, Berenguera A, Mandomando I, Sanz S, Aponte JJ, Sigauque B, Alonso PL, Menéndez C. Impact of maternal human immunodeficiency virus infection on birth outcomes and infant survival in rural mozambique. Am J Trop Med Hyg. 2009;80(5):870–6.
52. Stringer EM, Kendall MA, Lockman S, Campbell TB, Nielsen-Saines K, Sawe F, Cu-uvin S, Wu X, Currier JS. Pregnancy outcomes among hiv-infected women who conceived on antiretroviral therapy. PloS one. 2018;13(7):.
53. van Lettow M, Landes M, van Oosterhout J, Schouten E, Phiri H, Nkhoma E, Kalua T, Gupta S, Wadonda N, Jahn A, et al. Prevention of mother-to-child transmission of hiv: a cross-sectional study in malawi. Bull World Health Organ. 2018;96(4):256.
54. M.N.A.C. Malawi aids response progress report 2015. 2015.
55. Chimbwandira F, Mhango E, Makombe S, Midiani D, Mwansambo C, Njala J, Chirwa Z, Jahn A, Schouten E, Phelps BR, et al. Impact of an innovative approach to prevent mother-to-child transmission of hiv—malawi, july 2011–september 2012. MMWR Morb Mortal Wkly Rep. 2013;62(8):148.
56. Jahn A, Harries AD, Schouten EJ, Libamba E, Ford N, Maher D, Chimbwandira F. Scaling-up antiretroviral therapy in malawi. Bull World Health Organ. 2016;94(10):772.
57. Corbett M. Support for plwhas in malawi. Field Exchange 25. 2005;25:31.
58. Chihana ML, Price A, Floyd S, Mboma S, Mvula H, Branson K, Saul J, Zaba B, French N, Crampin AC, et al. Maternal hiv status associated with under-five mortality in rural northern malawi: a prospective cohort study. J Acquir Immune Defic Syndr 1999. 2015;68(1):81.
59. Dowswell T, Carroli G, Duley L, Gates S, Gülmezoglu AM, Khan-Neelofur D, Piaggio G. Alternative versus standard packages of antenatal care for low-risk pregnancy. Cochrane Database Syst Rev. 2015;7.

Publisher's Note

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com

frontiers in Nutrition

ORIGINAL RESEARCH
published: 18 November 2021
doi: 10.3389/fnut.2021.714232

Evaluating the Effect of Appropriate Complementary Feeding Practices on Child Growth in Malawi Using Cross-Sectional Data: An Application of Propensity Score Matching

Halima S. Twabi[1*†], *Samuel O. M. Manda*[2,3†] *and Dylan S. Small*[4‡]

[1] *Department of Mathematical Sciences, University of Malawi, Zomba, Malawi,* [2] *Biostatistics Research Unit, South African Medical Research Council, Pretoria, South Africa,* [3] *Department of Statistics, University of Pretoria, Pretoria, South Africa,* [4] *Department of Statistics, University of Pennsylvania, Philadelphia, PA, United States*

OPEN ACCESS

Edited by:
Samiran Bisai,
Regional Medical Research Centre for Tribals (ICMR), India

Reviewed by:
Suman Chakrabarty,
West Bengal State University, India
Sadaruddin Biswas,
Sidho Kanho Birsha University, India

***Correspondence:**
Halima S. Twabi
htwabi@unima.ac.mw

†*These authors share first authorship*

‡*Senior author*

Specialty section:
This article was submitted to Nutrition Methodology, a section of the journal Frontiers in Nutrition

Received: *05 August 2021*
Accepted: *28 September 2021*
Published: *18 November 2021*

Citation:
Twabi HS, Manda SOM and Small DS (2021) Evaluating the Effect of Appropriate Complementary Feeding Practices on Child Growth in Malawi Using Cross-Sectional Data: An Application of Propensity Score Matching. Front. Nutr. 8:714232. doi: 10.3389/fnut.2021.714232

Introduction: Appropriate complementary foods have been found to provide infants and young children with nutritional needs for their growth and development. In the absence of a randomized control trial (RCT), this study used observational data to evaluate the effect of appropriate complementary feeding practices on the nutritional status of children aged 6–23 months in Malawi using a propensity score matching statistical technique.

Methods: Data on 4,722 children aged 6 to 23 months from the 2015–16 Malawi Demographic and Health Survey (MDHS) were analyzed. Appropriate complementary feeding practices were assessed using the core indicators recommended by the World Health Organization (WHO)/United Nations Children's Fund (UNICEF), and consist of the introduction of complementary feeding, minimum dietary diversity, minimum meal frequency and minimum acceptable diet based on a dietary intake during a most recent 24-h period.

Results: The prevalence of stunting (height-for-age *z-score* < -2 SD) was 31.9% (95% CI: 29.3%, 34.6%), wasting (weight-for-height *z-score* < -2 SD) 3.5% (95% CI: 2.6%, 4.7%) and underweight (weight-for-age *z-score* < -2 SD) 9.9% (95% CI: 8.4%, 11.8%). Of the 4,722 children, 7.7% (95% CI: 6.9%, 8.5%) were provided appropriate complementary foods. Appropriate complementary feeding practices were found to result in significant decrease in stunting (OR = 0.7, 95% CI: 0.4, 0.95). They also resulted in the decrease of wasting (OR = 0.4, 95% CI: 0.1, 1.7) and underweight (OR = 0.6, 95% CI: 0.2, 1.7).

Conclusion: Appropriate complementary feeding practices resulted in a reduction of stunting, wasting, and underweight among children 6 to 23 months of age in Malawi. We recommend the continued provision of appropriate complementary foods to infants and young children to ensure that the diet has adequate nutritional needs for their healthy growth.

Keywords: average treatment effect, child growth and nutrition, complementary feeding practices, Sub-Saharan Africa, propensity score matching, quasi-experimental method

INTRODUCTION

Undernourished children have their immune system compromised, making them more vulnerable to chronic diseases and infections. They take longer to recover from illnesses and have an increased risk of dying, especially from infectious diseases (1). Malnutrition levels among children aged under 5 years are high in Sub-Saharan Africa (SSA). They are too short for their age (stunted) (32%), too thin for their height (wasted) (10%), and too thin for their age (underweight) (15%) (2–4). To enhance their nutrition needs, complementary iron, minerals and vitamin-rich foods are introduced at the age of 6 months. Complementary feeding entails giving a child other semi-solid, soft foods and solid foods, liquids, water along with breast milk (5). Several randomized control trials (RCTs) have shown the beneficial nutritional effect of optimal timing and amount of complementary foods on child growth (6–9). Similar findings of the effect of complementary foods were also found using observational studies (10–12).

Due to high levels of poverty, inadequate diversity in farming and foods, and high levels of HIV, under-nutrition among children in Malawi has remained a public health concern. Children in Malawi experience chronic under-nutrition and high levels of deficiencies in important micro-nutrients that are needed for optimal growth (13, 14). Thus, the Government of Malawi has initiated several policies, a Food and Nutrition Security Policy (2007) and a National Nutrition Policy and Strategic Plan (NNPSP) (15, 16). The Young Child Nutrition Strategy (2009–2014) (17, 18) set a number of priority areas, one of which was improved infant and young child feeding practices, with the aim of improving the intake of essential micro-nutrients. In particular, children aged 6–23 months are fed porridge made from Maize flour (grains) with the addition of legumes, such as groundnuts and soybeans, cooking oil and leafy vegetables to the porridge (19) for improved nutrition. There is a need to provide empirically driven evidence that evaluates the impact of child nutrition interventions in Malawi. However, there has been a paucity of randomized control studies that have evaluated and confirmed the beneficial effect of appropriate complementary feeding practices on child growth indicators.

Even though a randomized control trial (RCT) could be conceived, it would be very unethical based on evidence found in previous RCT studies on the beneficial effect of complementary feeding. However, we could still assess the efficacy of appropriate complementary feeding practices on growth measures of children 6–23 months of age using observational data with appropriate statistical techniques. In this study, we used observational complementary feeding data from the 2015–16 Malawi Demographic and Health Survey (2016 MDHS). The DHS Program provides very revealing health data on which most of the health indicators in the sub-Saharan African region are derived (20).

Our proposed method relies on balancing confounder variables between children who were provided appropriate complementary foods and those who were not. For example, a mother's education, counseling on child feeding, household wealth, a child's age, duration of breastfeeding, source of drinking water, media exposure, mother's body mass index, a mother's HIV status, previous history of infections, sex of a child and low birth weight are associated with child growth indicators (11, 21–25). A statistical method that could be used to balance the distribution of confounder variables is based on the Propensity Score (PS) methods (26–28). In this way, we would have derived a quasi-experiment to estimate the effect of appropriate complementary feeding practices on growth measurements of children 6–23 months of age based on the survey data.

MATERIALS AND METHODS

Study Data and Design

The data was obtained from the 2015–16 Malawian Demographic Health Survey (MDHS), which was implemented by the National Statistical Office of Malawi in conjunction with the Ministry of Health and the Community Health Services Unit of Malawi. More details on the sampling procedure and design can be obtained from the 2015–16 MDHS report (29). The study analyzed data on 4,722 children aged 6 to 23 months for whom weight (in kilograms) and height (in centimeters) were measured.

Outcome and Confounder Variables

We used three anthropometric indices for children as outcome variables, namely height-for-age, weight-for-height and weight-for-age. For our study, we followed the convention of recording each index as a *z-score*, which indicates how many standard deviations (SD) is a child's index from the median of the World Health Organization (WHO) Child Growth Standards (30). For example, height-for-age *z-score* < -2 SD denotes a child who is stunted; similarly for wasting (weight-for-height *z-score* < -2 SD), and underweight (weight-for-age *z-score* < -2 SD).

Covariates that were identified within the Malawi Demographic and Health Survey as potential confounders included: mother's education, counseling on breastfeeding, history of diarrhea, household wealth, child's age, mother's employment, sex of a child, place of residence, wealth quintile, mother's age, mother's HIV status, birth weight and antenatal visits. However, we replaced mother's employment status with mother's education and wealth as only a few mothers were employed full-time, hence, the matched sample would have been small within mothers belonging in the full-time category.

Measurement of Complementary Feeding Indicators

Appropriate complementary feeding practices were defined using the core indicators recommended by the WHO/UNICEF in 2008 (31) which combines the introduction of complementary feeding, minimum dietary diversity, minimum meal frequency and minimum acceptable diet based on a dietary intake 24-h before the survey and calculated for the age range 6–23 months of age. These are defined as:

a) Introduction of complementary foods = 1 if a child aged 6–23 months was complementary fed (solid, semi-solid or soft foods) and 0 otherwise (21, 31).
b) Minimum dietary diversity (MDD) = 1 if a child received foods from four or more food groups during the previous day and 0 otherwise. This refers to the child receiving the following food groups; grains, roots and tubers; legumes and nuts; dairy products (milk, yogurt and cheese); flesh foods (meat, fish, poultry and liver/organ meats); eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables (21, 31).
c) Minimum meal frequency (MMF) = 1 if a breastfeeding and non-breastfeeding child aged 6–23 months received complementary foods the minimum number of times or more (minimum was defined as two times for breastfed infants 6–8 months; three times for breastfed children 9–23 months; and four times for non-breastfed children 6–23 months) in the previous day (21, 31), 0 otherwise.
d) Minimum acceptable diet (MAD) = 1 if a child was fed a minimum dietary diversity and minimum meal frequency during the day or night preceding the survey (31) and 0 otherwise. Minimum acceptable diet was calculated as a composite indicator from the following;
    i Breastfed children—minimum dietary diversity and minimum meal frequency as above.
    ii Non-breastfed children—minimum dietary diversity but excluding the dairy products category (4 out of 6 groups) and minimum meal frequency and 2 or more milk feeds.

In measuring the appropriate complementary feeding practice exposure using the MDHS data, we set it equal to 1 if a child was provided complementary foods and had a minimum dietary diversity and a minimum meal frequency, otherwise, it was set to 0.

Statistical Analysis

The propensity score matching (PSM) was used to balance the covariate differences between children who were provided appropriate complementary foods and those who were not. In the estimation of propensity scores (PS), we used a binary logistic regression to model whether a child was provided appropriate complementary foods as a function of potential confounders. From this model, probabilities of whether a child was given appropriate complementary foods conditional on the confounders were estimated to give "propensity scores." The scores were then subsequently used in covariate matching between the two appropriate complementary feeding practice groups (26–28, 32, 33). We used the nearest neighbor method on a 1:1 ratio (34) at a recommended 0.2 caliper (35).

Covariate Balance Assessment

The standardized difference was used to assess balance for the distribution of measured potentially confounded covariates between appropriately and inappropriately complementary fed children (36). A standardized difference of 10% (0.1) is commonly used as a cut-off point to assess adequate balance of the covariates (36, 37). We further compare the balance in the measured confounders between children who were provided appropriate complementary foods and those who were not using a McNemar's Bowker's test for categorical variables and a paired *t*-test for continuous variables.

Sensitivity Analysis

One strong assumption of conducting the PSM is that there remains no unobserved confounding. It is impossible to prove that no unobserved confounding exists, but through a sensitivity analysis, we can measure if the results are sensitive to hidden bias. This was done by using a McNemar's exact test of sensitivity for binary outcomes. McNemar's test compares the number of discordant pairs in which appropriately complementary fed children had improved growth against inappropriately complementary fed children who did not have improved growth (38–40). Several packages are available to conduct a sensitivity test on the effects for binary outcomes in R (41) and Stata (42). We used the *binarysens* of *Rbound* package in R to obtain the values of the upper and lower bounds on the estimates for wasting, stunting and underweight (43).

RESULTS

Summary information about characteristics of feeding practices and prevalence of the growth indicators for the 4,722 studied children aged 6–23 months is presented in **Table 1**. Of these children, 3.7% (95% CI: 2.7%, 4.6%), 27% (24.8%, 29.3%), and 13.6% (95% CI: 12.1%, 15%) were wasted, stunted, and underweight, respectively. Of the 4,722 children, 3,725 (78.9%) were still breastfeeding while 997 (21.1%) were not breastfeeding at the time of the survey. A majority of the children 3,929 (83.2%) were given complementary foods, and of these children, 1,589 (42.7%) were aged 6–11 months. More than half of all the children (57.5 %) were fed more than two times a day, the day preceding the survey. The common dietary foods given to the children were grains (27.9%) followed by legumes and nuts (24.5%). Out of the 4,722 children, 3,942 (83.6%) children had mothers who were counseled on breastfeeding.

Table 2 presents the proportion of children meeting the four WHO/UNICEF feeding indicators. Of the 4,722 children, 1,434 (30.4%) children aged 6–23 months received solid, semi-solid or soft foods the minimum number of times meeting a minimum meal frequency (MMF), 1,073 (22.7%) children were offered four or more food groups on the day preceding the survey meeting the minimum dietary diet (MDD), and 342 (8.0%) children were given a minimum adequate diet. Combining the three feeding indicators, introduction of complementary foods, minimum dietary diversity and minimum meal frequency, the overall prevalence of children provided with appropriate complementary foods was 7.7 % (95% CI: 6.9, 8.5).

Propensity Score Estimation

We modeled the probability of whether a child was provided appropriate complementary foods as a function of potential confounders namely, mother's HIV status, mother's age, child's

TABLE 1 | Characteristics of complementary feeding and growth indicators among the 4,722 children 6–23 months old from the 2015–16 MDHS.

	Freq.	Percent	(95% Conf. Interval)
Growth indicators			
Wasted	65	3.5%	(2.6, 4.7)
Stunted	498	31.9%	(29.3, 34.6)
Underweight	168	9.9%	(8.4, 11.8)
Characteristics of child feeding			
Ever breastfed			
Yes	4,680	99.1%	(98.8, 99.4)
No	42	0.9%	(0.6, 1.2)
Still breastfeeding			
Yes	3,725	78.9%	(77.7, 80.1)
No	997	21.1%	(19.9, 22.3)
Current age of child still breastfeeding			
6–11 months	1,589	42.6%	(41.1, 44.2)
12–17 months	1188	31.9%	(30.4, 33.4)
18–23 months	948	25.5%	(24.1, 26.8)
Current age of child not being breastfed			
6–11 months	327	32.8%	(29.9, 35.7)
12–17 months	274	27.5%	(30.4, 33.4)
18–23 months	376	39.7%	(36.7, 42.8)
Counseled on breastfeeding			
Yes	3942	83.5%	(82.5, 84.6)
No	762	16.0%	(15.1, 17.2)
Don't know	20	0.5%	(0.2, 0.9)
Ever started complementary feeding			
Yes	3,929	83.2%	(82.1, 84.3)
No	793	16.8%	(15.7, 17.9)
Current age of the child on complementary feeding			
6–11 months	1,586	40.4%	(38.8, 41.9)
12–17 months	1,273	32.4%	(30.9, 33.9)
18–23 months	1,070	27.2%	(25.8, 28.6)
No. of times a child aged 6–23 months was fed			
Once only	1,440	30.5%	(29.2, 31.8)
2–3 times	2,715	57.5%	(56.1, 58.9)
4 + times	567	12%	(11.1, 12.9)
Type of dietary food			
Diary	922	19.5%	(18.4, 20.7)
Grain	1,318	27.9%	(26.6, 29.2)
Legumes and nuts	1,155	24.5%	(23.2, 25.7)
Meat	639	13.5%	(12.6, 14.5)
Eggs	409	8.7%	(7.9, 9.5)
VA rich fruits and vegetables	208	4.4%	(3.8, 4.9)
Other rich fruits and vegetables	71	1.5	(1.2, 1.9)

VA, Vitamin A.

TABLE 2 | Proportion of infants and young children meeting the four WHO/UNICEF feeding indicators in 2015–16.

Indicator	Frequency	Percent (95% confidence interval)
Minimum acceptable diet		
Yes	378	8.0% (7.2, 8.8)
No	4,344	92.0% (91.2, 92.8)
Minimum meal frequency		
Yes	1,434	30.4% (29.1, 31.7)
No	3,288	69.6% (68.3, 70.9)
Minimum dietary diversity		
Yes	1,073	22.7% (21.5, 23.9)
No	3,649	77.3% (76.1, 78.5)
Introduction of complementary foods		
Yes	784	16.6% (20.2, 24.5)
No	3,938	84.3% (82.3, 84.5)

Freq., frequency.

age, child's sex, household wealth, mother's education, antenatal care visits, place of residence, birth weight, history of diarrhea and the results are presented in **Table 3**. In the univariate associations, there was a significant association between older child age (12–17 months) (OR = 1.32, 95% CI: 1.01, 1.73), (18–23 months) (OR = 1.46, 95% CI: 1.12, 1.92), rural residence (OR = 0.44, 95 % CI: 0.3, 0.6), Secondary/Post-Secondary education for mothers (OR = 2.9, 95 % CI: 1.5, 5.6) and low birth weight (OR = 1.42, 95% CI: 1.04, 1.93) with the provision of appropriate complementary foods. However, in the multivariate logistic regression analysis, only the age of the child was associated with appropriate complementary feeding practice (as expected older children were more likely to be given complementary foods; for example, children aged 18–23 months were one and a half times more likely than children aged 6–11 months (OR = 1.68, 95% CI: 1.01, 2.78), (see **Table 3**).

The Hosmer and Lemeshow goodness-of-fit test was performed to check if the propensity score model fit the data well. The Hosmer and Lemeshow test statistic had a *p*-value of 0.5881 indicating that the model used to estimate the propensity scores fit the data reasonably well. In addition, a classification accuracy analysis test was done to check the percentage of children correctly specified as being provided with appropriate complementary foods. The results for classification analysis showed that 91.05% of the children were correctly specified to either being appropriately complementary fed or not.

Distribution of Confounders and Balance Assessment

Table 4 presents the distribution of the studied confounder variables before and after matching. We used the nearest neighbor matching algorithm with a 0.2 caliper to restrict the difference in propensity scores between matched children. Before matching on the propensity score, out of the 4,722 children, 368 (7.7%) were provided appropriate complementary foods and 4,354 (92.3%) were not provided appropriate complementary foods. Of the 368 children, 43.8% were born to HIV-uninfected mothers, while 56.2% were born to HIV-infected mothers. Of those children who were provided appropriate complementary foods, 205 (55.7 %) were males and 163 (44.3%) were females, 232 (63.1%) resided in the rural area and 136 (36.9%) were from the urban area, 325 (88.3%) had a normal weight and 43 (11.7 %) had a low birth weight, and 194 (52.7%) had mothers who had primary education. Place of residence, sex of a child, household

TABLE 3 | Odds ratio (95% Confidence Interval) for the association between provision of appropriate complementary foods and predictors of child growth.

Variable	Frequency	Univariate odds ratio (95% conf. interval)	Adjusted odds ratio (95% conf. interval)
Residence			
Urban	755 (15.9%)	1.00	1.00
Rural	3,967 (84.1%)	0.44 (0.32, 0.63)	0.82 (0.45, 1.46)
Mother's HIV status			
HIV negative	1,460 (30.9%)	1.00	1.00
HIV positive	127 (8.0%)	1.54 (0.81, 2.9)	1.4 (0.74, 2.65)
Mother's education			
None	524 (11.1%)	1.00	1.00
Primary	3,176 (67.3%)	1.69 (0.95, 3.04)	2.7 (1.12, 6.3)
	1,022 (21.6 %)	2.9 (1.54, 5.63)	3.0 (1.2, 7.9)
Secondary/Post-secondary			
Had diarrhea last 2 weeks			
No	2,960 (62.7%)	1.00	1.00
Yes	1762 (37.3%)	1.19 (0.89, 1.57)	1.12 (0.74, 1.68)
Wealth			
Poor	2,168 (45.9%)	1.00	1.00
Medium	903 (19.2%)	1.37 (0.93, 2.02)	0.66 (0.34, 1.26)
Rich	1,651 (34.9%)	2.43 (1.77, 3.34)	1.59 (0.96, 2.63)
Mother's age			
15–24 years	2,102(44.5%)	1.00	1.00
25–34 years	1,881 (39.8%)	1.1 (0.81, 1.44)	1.34 (0.88, 2.05)
35–49 years	739 (15.7%)	0.75 (0.5, 1.13)	0.92 (0.47, 1.80)
Child's age			
6–11 months	1,596 (33.8%)	1.00	1.00
12–17 months	1,648 (34.9%)	1.32 (1.01, 1.73)	1.38 (0.84, 2.27)
18–23 months	1,478 (31.3%)	1.46 (1.12, 1.92)	1.68 (1.01, 2.78)
Antenatal visits			
0–1	166 (3.5%)	1.00	1.00
2–3	2,195 (47%)	1.55 (0.52, 4.67)	1.17 (0.23, 5.99)
4 above	2361 (49.5%)	1.88 (0.60, 5.896)	1.33 (0.25, 7.10)
Birth weight			
Normal weight	4,257 (90.2%)	1.00	1.00
Low birth weight	465 (9.8%)	1.42 (1.04, 1.93)	1.08 (0.55, 2.11)
Sex of child			
Male	2,434 (51.5%)	1.00	1.00
Female	2,288 (48.5%)	1.31 (1.06, 1.62)	0.98 (0.68, 1.4)

Conf. Int, Confidence Interval.

wealth index, mother's educational status, and mother's age were distributed differently between children provided appropriate complementary foods and those not provided appropriate complementary foods ($P < 0.05$), (see **Table 4**).

We assessed for balance on the matched sample by checking that each confounder did not significantly differ in proportion between children who were given appropriate complementary foods and those who were not using McNemar's Bowker's test for categorical variables. After PS matching on the confounders, 368 matched pairs (appropriately and not appropriately complementary fed pairs of children) were analyzed for the difference in baseline characteristics as shown in **Table 4**. No significant difference was observed in the covariates between children provided appropriate complementary foods and those who were not provided.

Covariate Balance

A standardized difference of 0.1 (10 per cent) was used as a decision criterion for balance in the measured covariates (37). We observed homogeneity in the distribution of the covariates between children provided with appropriate complementary foods and those who were not ($P > 0.05$) on the matched sample (**Table 5**). Before matching, the absolute value of the standardized difference was >0.1 for mother's education, child's age, mother's age, sex of a child, and history of diarrhea. After matching, all covariates had an absolute standardized difference of <0.1 indicating a balance in differences within the observed covariates between children who were provided appropriately complementary foods and those who were not.

Estimation of Appropriate Complementary Feeding Effect

For the unmatched sample, an ordinary logistic regression was used to assess the effect of appropriate complementary feeding practices on wasting, stunting and underweight, accounting for potential confounders. For the effect estimation of appropriate complementary feeding practices on child growth on the matched data, conditional logistic regression models were used, the results are shown in **Table 6**. For the unmatched sample, there was no significant association between appropriate complementary foods and child growth even though there was an indication that appropriate complementary foods were protective. For the matched sample, appropriate complementary foods were associated with a reduction in stunting (OR = 0.7, 95% CI: 0.4, 0.95); wasting (OR = 0.4, 95% CI: 0.1, 1.7) and underweight (OR = 0.6, 95% CI: 0.2, 1.7), but the association was not significant for wasting and underweight.

Sensitivity Analysis

For a sensitivity analysis, we consider a range of values of the odds ratio of two matched children to be given appropriate complementary foods [this is denoted as gamma in the literature (40, 44)] from 1 to a maximum value of 2 with an increment of 0.1. If there were no unmeasured confounding, this maximum odds ratio would be 1; higher values of the maximum odds ratio correspond to more unmeasured confounding. The results for the sensitivity analysis for the upper bound *p*-values of McNemar's sensitivity test are presented in Table A1 in the **Appendix**. The finding of a positive effect of appropriate complementary foods on stunting was least robust to possible unobserved confounders. The significant level of gamma at which we would have to question our conclusion of a positive effect on stunting was very close to 1 (between 1.1 and 1.2). For wasting, it would require gamma values between 1.7 and 1.8 to render the conclusion of the beneficial effect of appropriate complementary foods on wasting spurious. In the case of the effect of appropriate

TABLE 4 | Distribution of characteristics between appropriately and inappropriately fed children before and after matching.

	Before matching			After matching		
	Appropriate ($n = 368$)	Not appropriate ($n = 4,354$)		Appropriate ($n = 368$)	Not appropriate ($n = 368$)	
Characteristic	***n* (%)**	***n* (%)**	***p*-value**	***n* (%)**	***n* (%)**	***p*-value***
Mother's HIV status						
HIV-uninfected	161 (43.8)	1,794 (41.2)		161 (43.8)	162 (44.0)	
HIV-infected	207 (56.2)	2,560 (58.8)	0.038	207 (56.2)	206 (56)	0.9068
Age of child						
6–11 months	114 (30.9)	1,485 (34.1)		114 (31)	118 (32.1)	
12–17 months	131 (35.6)	1,518 (34.9)		131 (35.6)	130 (35.3)	
18–23 months	123 (33.5)	1,351 (31.0)		123 (33.4)	120 (32.6)	
6–23 months	368 (100)	4,354 (100)	0.231	368 (100)	368 (100)	0.1624
Sex of child						
Male	205 (55.7)	2,467 (56.7)		205 (55.7)	213 (57.9)	
Female	163 (44.3)	1,887 (43.3)	0.05	163 (44.3)	155 (42.1)	0.2864
Residence						
Urban	136 (36.9)	1,028 (23.6)		136 (37)	137 (37.2)	
Rural	232 (63.1)	3,326 (76.4)	<0.001	232 (63)	231 (62.8)	0.6617
Mother's education						
None	45 (12.2)	811 (18.6)		45 (12.2)	39 (10.6)	
Primary	194 (52.7)	2,745 (63.1)		194 (52.7)	204 (55.4)	
Secondary & Post-Secondary	129 (35.1)	798 (18.3)	<0.001	129 (35.1)	125 (34)	0.6123
History of diarrhea						
Yes	207 (56.2)	2,443 (56.1)		207 (56.2)	207 (56.3)	
No	161 (43.8)	1,911 (43.9)	0.228	161 (43.8)	161 (43.7)	0.1564
Birth weight						
Normal weight	325 (88.3)	3,917 (89.9)		325 (88.3)	324 (88.0)	
Low birth weight	43 (11.7)	437 (10.1)	0.05	43 (11.7)	44 (12.0)	0.2164
Wealth						
Poor	101 (27.5)	2,075 (47.7)		101 (27.5)	100 (27.2)	
Medium	61 (16.6)	827 (18.9)		61 (16.6)	66 (17.9)	
Rich	206 (55.9)	1,452 (33.4)	<0.001	206 (55.9)	202 (54.9)	0.3369
Mother's age						
15–24 years	156 (42.4)	1,940 (44.6)		156 (42.3)	167 (45.4)	
25–34 years	161 (43.8)	1,726 (40.6)		161 (43.8)	149 (40.5)	
35–49 years	51 (13.9)	926 (15.8)	0.005	51 (13.9)	52 (14.1)	0.9728

*i) *p-values calculated using McNemar's test for dichotomous variables and Marginal homogeneity test for covariates with more than two categories for n = 368.*
ii) n, Number (frequency).
iii) p, p-value.
iv) HIV, Human Immune Deficiency Virus.

complementary foods on underweight, the effect was sensitive to an unobserved confounder at gamma values between 1.5 and 1.6.

DISCUSSION

The study set out to estimate the beneficial effect of appropriate complementary foods on anthropometric indices of children 6–23 months of age in Malawi using observational data from the 2015–16 Malawi Demographic and Health Survey. We applied propensity score matching to balance confounders between appropriate complementary feeding practice groups. We found that children who were provided appropriate complementary foods resulted in a decrease in stunting, wasting and underweight. To the best of knowledge, this is first study that has estimated the effect of appropriate complementary feeding practices on child growth using observational data by employing quasi-experimental statistical methods such as propensity score matching.

Similar findings were obtained using randomized control trials (11, 45–48), and (11, 49–51) using observational studies. Our findings confirm WHO guidelines (5, 52) that recommend initiating complementary feeding at 6 months of age while continuing to breastfeed, with adequate timing of meals and the appropriate diversity of the foods. In addition, the findings

TABLE 5 | Absolute standard differences and *p*-values before and after application of propensity matching.

	Before		After	
Confounder	**Abs. Std Diff**	***p*-value**	**Abs. Std Diff**	***p*-value**
Sex of a child	0.168	0.07	0.08	0.957
Mother's education	0.16	0.029	0.05	0.513
Child's age	1.22	0.003	0.003	0.426
Residence	0.06	0.11	0.04	0.825
Had diarrhea	0.27	0.034	0.04	0.179
Region	0.01	0.784	0.05	0.303
Mother's employment	0.04	0.409	0.06	0.197
Wealth	0.144	0.114	0.01	0.592
Mother's age	0.18	<0.001	0.01	0.417
Birth weight	0.003	0.155	0.04	0.064
Antenatal visit	0.05	0.75	0.00	0.486

Abs. std. diff, Absolute standardized differences.

TABLE 6 | The effect of appropriate complementary feeding practices on child growth indicators for children aged 6–23 months in Malawi.

Indicators	**Before matching**	**After matching**
	AOR (95% conf. interval)	**OR (95% conf. interval)**
Wasting	0.7 (0.2, 2.2)	0.4 (0.1, 1.7)
Stunting	0.8 (0.5, 1.3)	0.7 (0.4, 0.95)
Underweight	0.5 (0.2, 1.1)	(0.2, 1.7)

The odds ratio and 95% confidence interval were obtained by the conditional logistic regression analyses (n = 368).
ii) AOR, Adjusted odds ratio.
iii) OR, Odds ratio.
iv) Conf.Int, Confidence Interval.

support the importance of guidelines that emphasize appropriate feeding practices among young children in Malawi (53, 54).

Appropriate complementary feeding practices among children 6–23 months should be encouraged in Malawi. The timely introduction of appropriate complementary foods has been suggested by the World Health Organization (WHO) as an effective measure in promoting improved child growth (31). Strategies that promote appropriate feeding practices must be emphasized to ensure that children receive the appropriate diet, especially those living within poor households. Poor households in Malawi are known to be food insecure (19). Regular access to highly nutritious food such as meat, eggs, milk and cooking oil is a challenge, as a result, children may be given maize porridge only without additional supplements (9, 55). Child feeding interventions that improve complementary feeding practices by using locally available recipes should be encouraged and promoted.

One of the strengths of our study is that we used the propensity score (derived from a logistic regression) matching to balance confounders between children who were provided appropriate complementary foods and those who were not. Thus, we have avoided ethical concerns that could have arisen by allocating children into appropriately complementary fed and inappropriately complementary fed groups while available evidence shows a clear benefit of appropriate complementary feeding practices. The beneficial effect of the provision of appropriate complementary foods to children 6–23 months of age in this study has been ascertained using observational data, which is readily available and owned by the government of Malawi. Our analyses have shown the viability of using observational data to provide appropriate scientific evidence on the effect of a locally relevant intervention on a major public health issue.

There were some limitations in the study that could have affected our findings. For example, the use of cross-sectional data does not provide a temporal association between the provision of appropriate complementary foods and the growth indicators. Longitudinal studies allow for such temporal associations between exposures and outcomes to estimate beneficial effects. For our study, it could have been that some children were initiated on complementary feeding before the survey. Thus, some sort of temporal association could be assumed, implying that the estimated beneficial effects could be possible. Appropriate complementary feeding practice used as exposure in our study was self-reported by the mothers. Self-reporting is subject to recall bias and may have underestimated or overestimated the effects (56, 57). However, since the recall time was short (24 h before the survey), we assume that the bias could be minimal. Furthermore, as expected the sample size reduced from 4,722 to 736, which is one of the inherent problems of using the PS matching. We were limited to the confounder variables available in the data set used for our study. One of the unobserved confounders that were not measured from the data is exclusive breastfeeding (EBF). Children who may have been exclusively breastfed and then provided with appropriate complementary foods would have had better growth indicators than those who were not exclusively breastfed. However, we assume that the EBF status was already accounted for in both appropriate complementary feeding groups. Findings from the sensitivity analysis show that the observed effect on stunting could have been affected by the unobserved confounders, while those of wasting and underweight were not. Thus, the results from our study on the effect of appropriate complementary feeding practices on child growth must be interpreted with caution.

CONCLUSIONS

Appropriate complementary foods were found to result in the reduction of stunting, wasting and underweight among children 6 to 23 months of age in Malawi. We recommend continuation of nutritional interventions that promote improved feeding practices and intake of essential micro-nutrients for infants and young children as outlined in several Government of Malawi nutritional policies and strategies. These interventions will significantly aid in sustained reduction of retarded growth among children in Malawi and low-middle-income countries (LMIC).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found below: https://dhsprogram.com/data/available-datasets.cfm.

ETHICS STATEMENT

The MDHS study was ethically approved by the Malawi Health Research Committee, Institutional Review Board of ICF Macro, and Center for Disease and Control (CDC) in Atlanta, GA, USA, and Prevention IRB. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

HT performed data management, statistical analysis, and wrote the initial draft of the manuscript. SM conceived and suggested the direction for this paper, reviewed statistical analysis and results, and helped with the revision of the manuscript. DS helped with the concept and direction. All authors have read and approved the final manuscript.

FUNDING

HT's work on this manuscript was partially supported by the DELTAS Africa Sub-Saharan Africa Consortium for Advanced Biostatistics Training (SSACABT). (The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust and the UK Government.) HT was also supported by the L'Oreal-UNESCO for Women in Science Sub-Saharan Africa Young Talents Programme through the Young Talents award. She also thanks the South African Medical Research Council (SAMRC) for funding part of her research visit to the SAMRC. SOMM's research work on this study was supported by the SAMRC. The views expressed in this publication are those of the author(s) and not those of the funders (AAS, NEPAD Agency, Wellcome Trust or the UK Government, L'Oreal, SAMRC). The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

HT thanks the Biostatistics Research Unit of the SAMRC for hosting her PhD research visits. We acknowledge the DHS team for making the 2016 MDHS data available to us to use on this study.

REFERENCES

1. Rice AL, Sacco L, Hyder A, Black RE. Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries. *Bull World Health Organ.* (2000) 78:1207–21. doi: 10.1590/S0042-96862000001000007
2. WHO. *Global Targets 2025 to Improve Maternal, Infant and Young Children Nutrition*. Tech. Rep. World Health Organization. Geneva, Switzerland (2017).
3. UNICEF. *Children, Food and Nutrition: Growing Well in a Changing World*. New York, NY: UNICEF (2019).
4. UNICEF. *WHO, World Bank Group. Joint Child Malnutrition Estimates. Levels and Trends in Child Malnutrition, 2016*. New York, NY: UNICEF (2020).
5. WHO. *Infant and Young Child Feeding: Model Chapter for Textbooks for Medical Students and Allied Health Professionals*. Geneva: World Health Organization (2009).
6. Dewey KG, Adu-Afarwuah S. Systematic review of the efficacy and effectiveness of complementary feeding interventions in developing countries. *Matern Child Nutr.* (2008) 4:24–85. doi: 10.1111/j.1740-8709.2007.00124.x
7. Krebs NF, Hambidge KM, Mazariegos M, Westcott J, Goco N, Wright LL, et al. Complementary feeding: a global network cluster randomized controlled trial. *BMC Pediatr.* (2011) 11:4. doi: 10.1186/1471-2431-11-4
8. Martinez B, Webb MF, Gonzalez A, Douglas K, del Pilar Grazioso M, Rohloff P. Complementary feeding intervention on stunted Guatemalan children: a randomised controlled trial. *BMJ Paediatr Open.* (2018) 2:e000213. doi: 10.1136/bmjpo-2017-000213
9. Vazir S, Engle P, Balakrishna N, Griffiths PL, Johnson SL, Creed-Kanashiro H, et al. Cluster-randomized trial on complementary and responsive feeding education to caregivers found improved dietary intake, growth and development among rural Indian toddlers. *Matern Child Nutr.* (2013) 9:99–117. doi: 10.1111/j.1740-8709.2012.00413.x
10. Perkins JM, Jayatissa R, Subramanian S. Dietary diversity and anthropometric status and failure among infants and young children in Sri Lanka. *Nutrition.* (2018) 55:76–83. doi: 10.1016/j.nut.2018.03.049
11. Walters CN, Rakotomanana H, Komakech JJ, Stoecker BJ. Maternal determinants of optimal breastfeeding and complementary feeding and their association with child under-nutrition in Malawi (2015–2016). *BMC Public Health.* (2019) 19:1503. doi: 10.1186/s12889-019-7877-8
12. Pokharel P, Adhikari A, Lamsal P, Adhikari R. Effect of complementary feeding practices and nutritional status of children (6-23 months) in Tamang community, Ambhanjyang Vdc of Makwanpur. *Janaki Med Coll J Med Sci.* (2017) 5:22–32. doi: 10.3126/jmcjms.v5i1.17984
13. Lelijveld N, Seal A, Wells JC, Kirkby J, Opondo C, Chimwezi E, et al. Chronic disease outcomes after severe acute malnutrition in Malawian children (chrosam): a cohort study. *Lancet Glob Health.* (2016) 4:e654–62. doi: 10.1016/S2214-109X(16)30133-4
14. Kumchulesi G. Persistence of child malnutrition in Malawi: explanations from demographic and health surveys. *J Afr Dev Afr Finance Econ Assoc.* (2018) 20:69–75. Available online at: https://www.jstor.org/stable/10.5325/jafrideve.20.1.0069
15. Malawi Ministry of Agriculture M. *Food and Nutrition Security Policy*. Tech. Rep. Government of Malawi. Lilongwe, Malawi (2007).
16. Government of Malawi M. *National Nutrition Policy and Strategic Plan 2007 - 2012*. Tech. Report. Government of Malawi. Lilongwe, Malawi (2007).
17. Malawi Ministry of Health P. *Infant and Young Child Nutrition Strategy, 2009-2014*. Tech. Report. Malawi Ministry of Health and Population. Lilongwe, Malawi (2009).
18. Osendarp SJ, Shilpi F, Gondwe T, Pangapanga-Phiri I, Kalimbira A, Mtimuni B, et al. Determinants of *Reductions in Childhood Stunting in Malawi's Community-based Nutrition Programs Health, Nutrition, and Population Discussion Paper*. Washington DC: World Bank (2019). doi: 10.1596/32302
19. Geresomo N, Mbuthia E, Matofari J, Mwangwela A. Risk factors associated with stunting among infants and young children aged 6-23 months in Dedza district of central Malawi. *Afr J Food Agric Nutr Dev.* (2017) 17:12854–70. doi: 10.18697/ajfand.80.16730
20. Manda SO, Kandala N-B, Ghilagaber G. Advanced techniques for modelling maternal and child health in Africa. In: *Advanced Techniques for Modelling Maternal and Child Health in Africa*. Dordrecht: Springer (2014). p. 1–7. doi: 10.1007/978-94-007-6778-2_1
21. Kassa T, Meshesha B, Haji Y, Ebrahim J. Appropriate complementary feeding practices and associated factors among mothers of children age 6–23 months in southern Ethiopia, 2015. *BMC Pediatr.* (2016) 16:131. doi: 10.1186/s12887-016-0675-x

22. Nkoka O, Mhone TG, Ntenda PA. Factors associated with complementary feeding practices among children aged 6–23 months in Malawi: an analysis of the demographic and health survey 2015–2016. *Int Health.* (2018) 10:466–79. doi: 10.1093/inthealth/ihy047
23. Dhami MV, Ogbo FA, Osuagwu UL, Agho KE. Prevalence and factors associated with complementary feeding practices among children aged 6–23 months in India: a regional analysis. *BMC Public Health.* (2019) 19:1034. doi: 10.1186/s12889-019-7360-6
24. Li Z, Kim R, Vollmer S, Subramanian S. Factors associated with child stunting, wasting, and underweight in 35 low-and middle-income countries. *JAMA Netw Open.* (2020) 3:e203386–e203386. doi: 10.1001/jamanetworkopen.2020.3386
25. Twabi HS, Manda SO, and Small DS. Assessing the effects of maternal HIV infection on pregnancy outcomes using cross-sectional data in Malawi. *BMC Public Health.* (2020) 20:974. doi: 10.1186/s12889-020-09046-0
26. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika.* (1983) 70:41–55. doi: 10.1093/biomet/70.1.41
27. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics.* (1996) 52:249–64. doi: 10.2307/2533160
28. Austin PC. An introduction to propensity score methods for reducing the effect of confounding in observational. *Multivariate Behav Res.* (2011) 46:399–424. doi: 10.1080/00273171.2011.568786
29. National Statistical Office (NSO) [Malawi] and ICF. *Malawi Demographic and Health Survey 2015-16 Report*. Tech. Rep. National Statistics Office Malawi and ICF. Zomba, Malawi, and Rockville, Maryland, USA. NSO and ICF (2007).
30. WHO. *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. Geneva: World Health Organization (2006).
31. WHO. *Indicators for Assessing Infant and Young Child Feeding Practices: Part 2: Measurement*. Geneva: World Health Organization (2010).
32. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Stat Med.* (2004) 23:2937–60. doi: 10.1002/sim.1903
33. Cole S, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* (2008) 168:656–64. doi: 10.1093/aje/kwn164
34. Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-ami statin use. *Stat Med.* (2006) 25:2084–106. doi: 10.1002/sim.2328
35. Gu XS, Rosenbaum PR. Comparison of multivariate matching methods: structures, distances, and algorithms. *J Comput Graph Stat.* (1993) 2:405–20. doi: 10.1080/10618600.1993.10474623
36. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* (2009) 28:3083–107. doi: 10.1002/sim.3697
37. Normand S-LT, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol.* (2001) 54:387–98. doi: 10.1016/S0895-4356(00)00321-8
38. Hernan MA, Robins JM. *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC (2020).
39. Rosenbaum P. *Observational Studies*. 2nd ed. New York, NY: Springer-Verlag (2002). doi: 10.1007/978-1-4757-3692-2
40. Rosenbaum PR. Sensitivity analysis in observational studies. *Stat Ref Online.* (2005) 4:1809–14. doi: 10.1002/0470013192.bsa606
41. Rosenbaum PR, Small DS. An adaptive Mantel–Haenszel test for sensitivity analysis in observational studies. *Biometrics.* (2017) 73:422–30. doi: 10.1111/biom.12591
42. Becker SO, Caliendo M. Sensitivity analysis for average treatment effects. *Stata J.* (2007) 7:71–83. doi: 10.1177/1536867X0700700104
43. Keele L. *An Overview of Rbounds: An R Package for Rosenbaum Bounds Sensitivity Analysis With Matched Data*.Vol. 1. Columbus, OH: White Paper (2010). p. 15.
44. Aakvik A. Bounding a matching estimator: the case of a Norwegian training program. *Oxf Bull Econ Stat.* (2001) 63:115–43. doi: 10.1111/1468-0084.00211
45. Marriott BP, White A, Hadden L, Davies JC, Wallingford JC. World Health Organization (who) infant and young child feeding indicators: associations with growth measures in 14 low-income countries. *Matern Child Nutr.* (2012) 8:354–70. doi: 10.1111/j.1740-8709.2011.00380.x
46. Disha A, Rawat R, Subandoro A, Menon P. Infant and young child feeding (iycf) practices in Ethiopia and Zambia and their association with child nutrition: analysis of demographic and health survey data. *Afr J Food Agric Nutr Dev.* (2012) 12:5895–914. doi: 10.4314/AJFAND.V12I2
47. Udoh EE, Amodu OK. Complementary feeding practices among mothers and nutritional status of infants in Akpabuyo area, cross river state Nigeria. *Springerplus.* (2016) 5:2073. doi: 10.1186/s40064-016-3751-7
48. Lassi ZS, Rind F, Irfan O, Hadi R, Das JK, Bhutta ZA. Impact of infant and young child feeding (iycf) nutrition interventions on breastfeeding practices, growth and mortality in low-and middle-income countries: systematic review. *Nutrients.* (2020) 12:722. doi: 10.3390/nu12030722
49. Kumar D, Goel N, Mittal PC, Misra P. Influence of infant-feeding practices on nutritional status of under-five children. *Indian J Pediatr.* (2006) 73:417–21. doi: 10.1007/BF02758565
50. Zongrone A, Winskell K, Menon P. Infant and young child feeding practices and child undernutrition in Bangladesh: insights from nationally representative data. *Public Health Nutr.* (2012) 15:1697–704. doi: 10.1017/S1368980012001073
51. Lassi ZS, Das JK, Zahid G, Imdad A, Bhutta ZA. Impact of education and provision of complementary feeding on growth and morbidity in children less than 2 years of age in developing countries: a systematic review. *BMC Public Health.* (2013) 13:S13. doi: 10.1186/1471-2458-13-S3-S13
52. WHO. *Global Strategy for Infant and Young Child Feeding*. Geneva: World Health Organization (2003).
53. Malawi Ministry of Health P. *Infant and Young Child Nutrition Policy and Guidelines 2003 - 2020*. Tech. Rep. Malawi Ministry of Health and Population. Lilongwe, Malawi (2011).
54. Phiri FP. Experiences of multi-sector programming in Malawi. *Field Exchange.* (2016) 52:97. Available online at: www.ennonline.net/fex/52/multisectorprogrammingmalawi
55. Kinyoki DK, Kandala N-B, Manda SO, Krainski ET, Fuglstad G-A, Moloney GM, et al. Assessing comorbidity and correlates of wasting and stunting among children in Somalia using cross-sectional household surveys: 2007 to 2010. *BMJ Open.* (2016) 6:e009854. doi: 10.1136/bmjopen-2015-009854
56. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc.* (2016) 9:211. doi: 10.2147/JMDH.S104807
57. Paeratakul S, Popkin B, Kohlmeier L, Hertz-Picciotto I, Guo X, Edwards L. Measurement error in dietary data: implications for the epidemiologic study of the diet–disease relationship. *Eur J Clin Nutr.* (1998) 52:722–7. doi: 10.1038/sj.ejcn.1600633

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

TABLE A1 | Sensitivity analysis - Rosenbaums upper bounds for stunting, wasting and underweight.

Gamma	Upper bound probability	Stunting	Wasting	Underweight
1	0.50	0.0089	<0.001	<0.001
1.1	0.52	0.0421	<0.001	0.0004
1.2	0.55	0.1269	<0.001	0.0025
1.3	0.57	0.2729	0.0002	0.0087
1.4	0.58	0.4569	0.001	0.0225
1.5	0.60	0.6373	0.0047	0.0467
1.6	0.62	0.7821	0.0145	0.0829
1.7	0.63	0.8809	0.0365	0.1310
1.8	0.64	0.9401	0.0768	0.1894
1.9	0.66	0.9719	0.1395	0.2553
2	0.67	0.9876	0.2241	0.3256

Maternal and Child Health Journal (2022) 26:2346–2354
https://doi.org/10.1007/s10995-022-03462-x

Estimating Unhealthy Food Effects on Childhood Overweight in Malawi Using an Observational Study

Halima S. Twabi[1] · Samuel O. M Manda[2] · Dylan S. Small[3]

Accepted: 18 July 2022 / Published online: 10 August 2022

Abstract

Introduction Consumption of unhealthy foods in children contributes to high levels of childhood obesity globally. In developing countries there is paucity of empirical studies on the association. This study employed propensity-score methods to evaluate the effect of unhealthy foods on overweight among children in Malawi using observational data.

Methods Data on 4625 children aged 6 to 59 months from the 2015-16 Malawi Demographic and Health Survey (MDHS) were analyzed. A multivariable logistic regression model of unhealthy foods (yes or no) on purported confounders of childhood overweight was used to obtain a child's unhealthy food propensity score. The propensity scores were then used to form matched sets of healthy and unhealthy fed children. The association between unhealthy foods and childhood overweight was assessed using the conditional logistic regression model.

Results The prevalence of overweight (body mass index (BMI) z-score > 2 standard deviations) was estimated at 4.5% (3.8%, 5.3%). The proportion of children who consumed unhealthy foods was estimated at 14.6% (95% CI: 13.1%, 16.2%). Our propensity score matching achieved a balance in the distribution of the confounders between children in the healthy and unhealthy food groups. Children fed unhealthy foods were significantly more likely to be overweight than those fed healthy foods (OR = 2.5, 95% CI: (1.2, 5.2)).

Conclusion The findings suggest the adverse effects of unhealthy foods on childhood overweight in Malawi. Thus, efforts to reduce unhealthy food consumption among children should be implemented and supported to address the problem of childhood overweight in Malawi and the sub-Saharan African region.

Keywords Propensity score · Child feeding · BMI · Overweight · Unhealthy foods · Observational studies

Significance

Childhood overweight is an emerging public health issue in Sub-Saharan Africa. Sustaining appropriate child feeding practices is a challenge in the region due to poverty and urbanization children are fed ready-made processed foods. There is scarcity of randomized controlled trials within the region that assess the impact of consumption of unhealthy foods on childhood overweight/obesity. Most data available on feeding practices have used observational studies. To determine the effect of consumption of unhealthy foods on childhood overweight, we use propensity-score methods that balance potential confounders for observational studies. Consumption of unhealthy foods had an adverse effect on childhood overweight.

Introduction

The increase in the consumption of unhealthy diets, such as salty, sugary, and saturated fatty foods, among children, has been associated with high childhood overweight/obesity

✉ Halima S. Twabi
htwabi@unima.ac.mw

Samuel O. M Manda
samuel.manda@up.ac.za

Dylan S. Small
dsmall@wharton.upenn.edu

[1] Department of Mathematical Sciences, University of Malawi, Zomba, Malawi

[2] Department of Statistics, University of Pretoria, Pretoria, South Africa

[3] Department of Statistics, University of Pennsylvania, Philadelphia, USA

levels in sub-Saharan Africa (SSA) (Kalimbira & Gondwe, 2015; Biadgilign et al., 2017;Tuyet et al., 2019). This has led to an increase in obesity-related health conditions such as type-2 diabetes mellitus, cardiovascular disease, asthma, and respiratory diseases within the region (Sahoo et al.,2015; Kansra et al., 2020). Globally, approximately 38 million under-five children were at least overweight/obese in 2018, and a quarter was in sub-Saharan Africa (Biadgilign et al., 2017; WHO, 2018). As regards child feeding practices, the World Health Organisation (WHO) recommends feeding children aged 6 to 24 months appropriate complementary foods (World Health Organization, 2008). This type of recommendation is often not sustained in low-income countries, for example, in SSA, where the children are often fed food that is commercially processed and has high energy content and low nutrients (Rousham et al., 2022).

Previous studies and reviews have highlighted the inconclusive and conflicting evidence on the effect of unhealthy food consumption on higher Body Mass Index (BMI) in children(English et al., 2019; Pries, Filteau, et al., 2019; Pries, Rehman, et al., 2019; Min et al., 2021; Rousham et al., 2022). Most of these studies have been conducted in high-income countries (English et al., 2019; Pries, Filteau, et al., 2019) . Ascertaining a true effect of unhealthy foods on childhood overweight would require a randomized controlled trial (RCT); however, this could be subject to ethical issues. A recent systematic review and meta-analysis found only two studies that were RCT (Rousham et al., 2022). In addition, for most of these observational studies, the analysis was done using standard statistical techniques. Observational studies could be challenging for evaluating effects due to the problem of confounder imbalance between unhealthy and healthy food groups. Statistical methods that are used to balance the confounders between the two food groups for these observational studies have been developed (Austin, 2011; Rosenbaum & Rubin, 1983; Rubin & Thomas, 1996). In particular, propensity-score based methods can be used to estimate the effect of treatment exposure (unhealthy foods) on the outcome (childhood overweight). This study aimed to use the Malawi Demographic and Health Survey (MDHS) data to assess the effect of unhealthy food diets on childhood overweight using a covariate adjustment method based on the propensity score.

Materials and methods

Study Data and Design

The data was obtained from the 2015-16 Malawi Demographic Health Survey (MDHS), implemented by the National Statistical Office of Malawi in conjunction with the Ministry of Health and Malawi's Community Health Services Unit. The sampling procedure and design can be obtained from the 2015-16 MDHS report (NSO, 2017). The study analyzed data on 4625 children aged 6 to 59 months for whom weight (in kilograms) and height (in centimeters) were measured as per described in the 2015-16 MDHS report.

Childhood overweight and associated confounders

A child was categorized as overweight if his/her weight by height (body mass index) z-score was above 2 standard deviations from the median of the reference child population (WHO, 2006). Our analysis excluded children who had missing or "Flagged cases" for the weight by height z-scores. Covariates identified within the MDHS as potential confounders were categorized as child or parent-related. These included the following: parent-related; maternal education, mother's age, antenatal visits, and mother's employment status. Child-related; history of diarrhea, whether a child had fever, household wealth, child's age, sex of a child, place of residence, and birth weight.

Child unhealthy foods

Following WHO/UNICEF (World Health Organization, 2008;World Health Organization, 2019), a child is fed unhealthy foods if their diet contains the following food items: sugar-sweetened foods and beverages such as juices, soft drinks (fizzy drinks), chocolates, sweets, pastries, and cakes; foods with oils, butter, saturated fats; processed foods such as cheese, yogurt, and other milk products.

Statistical methods

The propensity score was obtained as the probability of a child being fed unhealthy foods given the measured confounders. This was done by fitting a multivariable logistic regression on child unhealthy feeding (yes/no).We implemented the propensity score matching using one-to-one matching (Gu & Rosenbaum, 1993; Austin & Mamdani 2006).The effect of unhealthy foods on childhood overweight was estimated using a conditional logistic regression model on the matched sample.

Results

Summary information about characteristics of feeding practices and prevalence of overweight for the 4625 children aged 6–59 months is presented in Table 1. Out of the 4625 children, 4.5% (95% CI: (3.8%, 5.3)) were overweight. Of

Table 1 Characteristics of child feeding and overweight among the 5110 under-five children, 2015-16 MDHS

Characteristic	Freq.	Percent	(95% CI)
Overweight	203	4.5%	(3.8, 5.3)
Type of complementary			
Healthy	3861	84.4	(82.6, 85.9)
Unhealthy	764	15.6	(14.1, 17.4)
Type of food			
Diary	441	9.8	(8.6, 11.2)
Grain	557	12.1	(10.8, 13.5)
Legumes and nuts	497	10.0	(8.7, 11.5)
Meat	314	6.6	(5.7, 7.6)
Eggs	142	3.0	(2.4, 3.9)
Fruits and vegetables	74	1.4	(0.9, 2.3)
Sugar-sweetened food and beverages			
Yes	600	12.5	(11.1, 13.9,)
No	4025	87.5	(86.1, 88.9)
Fatty and oily foods			
Yes	243	4.9	(4.1, 5.9)
No	4382	95.1	(94.1, 96)

Freq – Frequency; 95% CI – 95% Confidence Interval

these, 764 children (15.6%) (95% CI: (14.1%, 17.4)) were fed unhealthy foods while 3861 (84.4%) were fed healthy foods. The common dietary food given to the children was grain (12.1%) followed by legumes and nuts (10.0%) and then diary (9.8%). Out of the 4625 children, 600 (12.5%) (95% CI: (11.1%, 13.9)) children had sugar-sweetened foods and beverages and 243 (4.9%) (95% CI: (4.1%, 5.9)) had saturated fatty and oily foods.

Propensity score estimation

We modelled the probability of unhealthy feeding by a child as a function of potential confounders namely, mother's age, child's age, child's sex, household wealth, mother's education, antenatal care visits, place of residence, birth weight, history of diarrhoea, mother's employment status and whether a child had fever. The results of the association between the confounders and log odds of eating unhealthy foods are presented in Table 2. In the univariate analysis, a child's age (25–59 months) (Coef.=-2.11, 95% CI: -2.3, -1.9), rural residence (Coef.=-0.8, 95% CI:-1.1,-0.4); a mother's secondary/post-secondary education (Coef.=0.7, 95% CI:0.2, 1.2); antenatal visits (2–3 times) (Coef.=0.8, 95% CI:0.5, 1.3), (4 visits and above) (Coef.=0.7, 95% CI:0.4, 1.2); children who had diarrhea (Coef.=0.6, 95% CI:0.4, 0.8), and children with mothers who were rich (Coef.=0.4, 95% CI: 0.2, 0.7) were associated with unhealthy feeding. For the multivariate analysis, children aged 25–59 months, children with mothers who had secondary or post-secondary education and those from rich families were more likely to have unhealthy foods. While children residing in rural areas were less likely to have been given unhealthy foods.

The Hosmer and Lemeshow goodness-of-fit test was performed to check if the propensity score model fit the data well. A p-value of 0.4719 was obtained indicating that the model fit the data reasonably well. Besides, a classification accuracy analysis test was done to check the percentage of children correctly specified into the unhealthy fed group. We observed that 92.1% of the children were correctly specified to either unhealthy feeding or not. The propensity of a child being fed unhealthy food was then calculated from the estimates obtained from the fitted multivariable logistic regression model.

Distribution of confounders and balance assessment

Table 3 presents the distribution of the studied confounder variables before and after matching. Before matching the children based on their propensity scores, we assessed the balance of the confounders between the feeding groups. Out of the 4625 children, 764 were fed unhealthy foods, and 3861 were fed healthy foods. Of the 764 children given unhealthy foods, 382 (48.9%) of the males were fed unhealthy foods, while 1893 (48.2%) males were fed healthy foods. Out of the 4625 children, 162 (21.4%) from the urban areas were fed unhealthy foods, while 11.3% were fed healthy foods. Most children aged 6–24 months were fed unhealthy foods (76.2%), while 28.3% were fed healthy foods. Of the children from rich households, 40.8% were fed unhealthy foods, and 32.3% had healthy foods. For children who had mothers with secondary or post-secondary education, 29.3% had unhealthy foods, and 19.3% had healthy foods. Place of residence, child's age, household wealth index, mothers' educational status, history of diarrhea, and mothers' age were distributed differently between children who had unhealthy and healthy foods ($P < 0.05$), see Table 3. We further assessed the systematic differences of the confounders using an absolute standardized difference. A standardized difference of 0.1 (10%) was used to denote meaningful balance between children given unhealthy and healthy foods across the measured confounders (Normand et al., 2001; Austin, 2009).The absolute value of the standardized difference of the covariates before matching was greater than 0.1 for mother's education, child's age, mother's age (35 years above), place of residence, mother's wealth status, antenatal visits (4 times above) and if a child had diarrhea.

We performed both the one-to-one (1:1) and the four-to-one (4:1) nearest neighbor matching algorithm with an acceptable difference between the scores of 0.2. A comparison of the matching algorithms between the groups was done and the 4:1 matching algorithm did not produce

Table 2 Association (on log-odds scale) between diet type and confounders of obesity, 2015-16 MDHS

Variable	Frequency	Univariate Coef. (95% CI)	Adjusted Coef. (95% CI)
Residence			
Urban	740 (12.8%)	ref	ref
Rural	3885 (87.2%)	-0.76(-1.1,-0.4)	-0.59 (-1.0, -0.1)
Mother's Education			
None	580 (13.1%)	ref	ref
Primary	3018 (66.0%)	0.18(-0.2,0.6)	-0.1 (-0.5, 0.4)
Secondary/Post-Secondary	1027 (20.9%)	0.70(0.2, 1.2)	0.36(-0.2, 0.9)
Had diarrhoea last 2 weeks			
No	3631 (78.5%)	ref	ref
Yes	994 (21.5%)	0.60(0.4,0.8)	0.2 (-0.1, 0.4)
Wealth			
Poor	2045 (47.2%)	ref	ref
Medium	904 (19.2%)	0.20(-0.1, 0.5)	0.18 (-0.2, 0.5)
Rich	1676 (33.6%)	0.44(0.2, 0.7)	0.36(0.1, 0.7)
Mothers age			
15–24 years	1636(36.6%)	1.00	ref
25–34 years	2103 (45.1%)	-0.3(-0.5, -0.02)	-0.1(-0.40, 0.2)
35–49 years	886 (18.3%)	-0.3(-0.6,0.1)	0.1(-0.3,0.5)
Child's age			
6–24 months	1682 (35.8%)	ref	ref
25–59 months	2943 (64.3%)	-2.1 (-2.3, -1.9)	1.95(1.6,2.4)
Antenatal visits			
0–1	137 (4.0%)	ref	ref
2–3	1603 (44.3%)	0.8(0.5,1.3)	0.1(-0.5, 0.7)
4 above	1868 (51.7%)	0.7(0.4,1.2)	0.1(-0.5, 0.7)
Birth weight			
Normal weight	4176 (90.2%)	ref	ref
Low birth weight	449 (9.8%)	-0.2(-0.6,0.2)	-0.04(-0.5,0.4)
Fever			
No	3209 (68.7%)	ref	ref
Yes	1416 (31.3%)	0.2(-0.03,0.4)	0.3(-0.01,0.5)
Mother's employment			
Not employed	1330 (28.0%)	ref	Ref
Employed	3295 (72.0%)	-0.1(-0.4,0.2)	0.07(-0.2,0.4)
Sex of child			
Male	2275 (48.3%)	ref	Ref
Female	2350 (51.7%)	-0.02(-0.2, 0.2)	0.05(-0.2,0.3)

Coef – Coefficient, 95% CI– 95% Confidence Interval

better balance between the scores as compared to the 1:1 matching. The analysis was done using the matched sample obtained using the 1:1 matching algorithm.

We assessed the balance on the matched sample by checking that there was no difference between the exposure groups across the observed confounders using a McNemar's Bowker's test for categorical variables (Bowker, 1948). After PS matching on the confounders, 584 children who had unhealthy foods were matched to 584 children given unhealthy foods. We analyzed the matched sample to assess the difference in baseline characteristics as shown in Table 3. No significant difference was observed in the covariates between children given unhealthy foods and those given healthy foods. Furthermore, the absolute standardized difference were less than 0.1 implying homogeneity in the distribution of all the covariates between children who had unhealthy foods and those who had healthy foods ($P > 0.05$) in the matched sample, see Table 4.

Estimation of unhealthy effect on childhood overweight

For the unmatched sample, the ordinary logistic regression model was used to assess the association between unhealthy feeding and childhood overweight, accounting for the potential confounders. As shown in Table 5, the effect of unhealthy foods on childhood overweight was not significant ((OR = 1.0, 95% CI: (0.6, 1.8)).

However, in using the propensity score matching methods, a child fed unhealthy foods was more likely to be

Table 3 Distribution of characteristics before and after matching

	Before Matching			After Matching		
	Unhealthy (n = 764)	**Healthy (n = 3861)**		**Unhealthy (n = 584)**	**Healthy (n = 584)**	
Characteristic	**N (%)**	**N (%)**	**p-value**	**N (%)**	**N (%)**	**p-value***
Age of child						
6–24 months	591 (76.2)	1091(28.3)		434(74.3)	416(71.2)	
25–59 months	173(23.8)	2770(71.7)	<0.001	150(25.7)	168(28.8)	0.483
Sex of child						
Male	382(48.9)	1893(48.2)		292(48.1)	292(49.6)	
Female	382 (51.1)	1968(51.8)	0.775	292(51.9)	292(50.4)	0.359
Residence						
Urban	162(21.4)	578(11.3)		135(22.5)	128(20.4)	
Rural	602(78.6)	3283(88.7)	< 0.001	449(77.5)	456(79.7)	0.882
Mothers' Education						
None	72(10.4)	508(13.8)		51(9.5)	38(6.3)	
Primary	472(60.3)	2546(66.9)		351(59.2)	365(64.1)	
Secondary & Post-Secondary	220(29.3)	807(19.3)	< 0.001	182(31.4)	181(29.7)	0.783
History of diarrhoea						
No	533(68.9)	3098(79.2)		362(60.6)	361(60.8)	
Yes	231(31.0)	763(20.8)	< 0.001	222(39.4)	223(39.2)	0.353
Birth weight						
Normal weight	700(91.8)	3476(89.9)		553(91.3)	544(93.9)	
Low birth weight	64(8.2)	385(10.1)	0.237	51(8.7)	40(6.1)	0.117
Wealth						
Poor	287(40.4)	1758(48.7)		209(39.3)	218(40.7)	
Medium	149(18.7)	755(19.1)		112(19.0)	114(17.0)	
Rich	328(40.8)	1348(32.3)	0.004	263(41.7)	252(42.3)	0.311
Had Fever						
No	503 (65.2)	2706 (69.3)		355 (59.0)	347 (57.7)	
Yes	261 (34.8)	1155 (30.7)	0.084	229 (41.0)	237 (42.3)	0.052
Mother's Employment status						
Not employed	218 (29.7)	1112(27.7)		168 (29.6)	168 (26.6)	
Employed	546 (70.3)	2749(72.3)	0.481	416 (70.4)	416 (73.4)	0.572
Maternal Age						
15–24 years	337(41.5)	1299(34.2)		269 (44.5)	269 (47.0)	
25–34 years	319(42.1)	1784(46.4)		235(40.1)	233 (39.7)	
35–49 years	108(16.4)	778(19.4)	0.036	80 (15.4)	82 (13.3)	0.071

i) * p–values calculated using McNemar's test for dichotomous variables and Marginal homogeneity test for covariates with more than two categories for n=275

ii) N – Number (frequency)

iii) p – p–value

overweight (OR = 2.5, 95% CI: (1.2, 5.2)) compared to those who had healthy foods. Children in urban areas who were fed unhealthy foods were significantly more likely to be overweight (OR = 3.4, 95% CI: (0.7, 16.7)). However, this was not the case for rural children (OR = 2.2, 95% CI: (1.0, 5.2)), (Table 6). A number of variables had missing data. We performed basic imputations and re-run the analysis. The estimated effects of unhealthy foods were hardly affected.

Sensitivity analysis

In the application of the propensity score matching, there is a need to consider the possible influence of unmeasured confounders on the estimated effects (Rosenbaum, 2005; Rosenbaum & Small, 2017). This was done using the Mantel Haenszel's sensitivity test (Aakvik, 2001; Rosenbaum, 2005)). We found that possibility of unmeasured confounders would not greatly affect the estimated effect of unhealthy foods on child overweight.

Breastfeeding duration is an important confounder of child feeding and growth (Martin, 2001; Haider & Saha, 2016; Lang Morović & Musić Milanović, 2019). Thus, we also re-run the analyses for only children with breastfeeding duration. There was no major differences in the estimation of the effect between unhealthy foods and childhood overweight.

Table 4 Absolute Standard Differences and p-values Before and After Matching

	Before		After	
Confounders	**Abs. Std. Difference**	**p-values**	**Abs. Std. Difference**	**p-values**
Sex of a child				
Male	0.019	0.624	0.000	1.000
Female	0.019	0.624	0.000	1.000
Mothers' education				
None	0.127	0.008	0.084	0.153
Primary	0.100	0.026	0.049	0.400
Secondary/ Post-Secondary	0.199	< 0.001	0.004	0.950
Child's age				
6–24 months	1.130	< 0.001	0.010	0.161
25–59 months	1.130	< 0.001	0.000	1.000
Residence				
Urban	0.176	< 0.001	0.029	0.624
Rural	0.176	< 0.001	0.029	0.624
Had diarrhoea				
No	0.341	< 0.001	0.004	0.952
Yes	0.341	< 0.001	0.004	0.952
Fever				
No	0.127	0.004	0.028	0.633
Yes	0.127	0.004	0.028	0.633
Wealth				
Poor	0.159	0.001	0.032	0.585
Medium	0.010	0.824	0.009	0.882
Rich	0.166	< 0.001	0.038	0.517
Mothers' age				
15–24 years	0.263	< 0.001	0.000	1.000
25–34 years	0.113	0.024	0.007	0.905
35 years above	0.190	< 0.001	0.010	0.866
Birth weight				
Low birth weight	0.026	0.174	0.070	0.231
Normal birth weight	0.026	0.174	0.070	0.231
Mother's Employment status				
Not employed	0.006	0.882	0.000	1.000
Employed	0.006	0.882	0.000	1.000
Antenatal visit				
0–1 times	0.009	0.845	0.206	0.257
2–3 times	0.002	0.961	0.047	0.724
4 above	0.006	0.902	0.035	0.953

Abs. Std. Diff – Absolute Standardized Differences

Table 5 Association between unhealthy feeding and overweight before matching

Characteristics	Overweight for height
	OR (95% CI)
Type of foods	
Healthy	1.00
Unhealthy	1.0 (0.6, 1.8)
Gender	
Male	1.00
Female	0.7 (0.5, 1.1)
Mothers' education	
None	1.00
Primary	0.8 (0.3, 2.1)
Secondary/Post-Secondary	0.5 (0.2, 1.5)
Child's age	
6–24 months	1.00
25–59 months	0.5 (0.3, 0.9)
Residence	
Urban	1.00
Rural	2.2 (1.1, 4.5)
Had diarrhoea	
No	1.00
Yes	0.4 (0.3, 0.8)
Fever	
No	1.00
Yes	0.6 (0.4, 1.0)
Wealth	
Poor	1.00
Medium	1.2 (0.7, 2.2)
Rich	1.6 (0.9, 2.7)
Mothers' age	
15–24 years	1.00
25–34 years	1.2 (0.7, 2.0)
35 years above	0.6 (0.3, 1.4)
Birth weight	
Low birth weight	1.00
Normal birth weight	1.5 (0.6, 4.0)
Mother's Employment status	
Not employed	1.00
Employed	1.0 (0.6, 1.5)
Antenatal visit	
0–1 times	1.00
2–3 times	0.5 (0.2, 1.1)
4 above	0.5 (0.2, 1.2)

Discussion

Childhood overweight is an emerging public health problem in SSA, where most children are given unhealthy food. Evidence on the impact of unhealthy foods on childhood overweight is conflicting due to the paucity of RCTs within the region. This study used observational data to determine the association between the consumption of unhealthy foods and childhood overweight in Malawi. The propensity score matching was used to balance the confounder distribution between children fed unhealthy foods and those fed healthy foods. Multivariate logistic regression was used to model a child's dietary food as a function of measured confounders to obtain the propensity of eating unhealthy foods using the 2015-16 MDHS. The results showed that children fed

Table 6 The effect of unhealthy foods on childhood obesity, Malawi 2015-16

Indicators	Overall	Urban	Rural
	OR (95% Conf. Int)	**OR (95% Conf. Int)**	**OR (95% Conf. Int)**
Overweight	2.5 (1.2, 5.2)	3.4 (0.7, 16.7)	2.2 (1.0, 5.2)

i) The odds ratio and 95% confidence interval were obtained by the conditional logistic regression analyses (n=368)

ii) AOR – Adjusted odds ratio

iii) Approp. – Appropriately complementary fed

iv) OR – Odds ratio

v) Conf.Int – Confidence Interval

unhealthy foods were more likely to be overweight than those provided healthy foods.

Several studies have also found a significant impact of consuming unhealthy snacks and sugar-sweetened beverages on childhood overweight (O'Connor et al., 2006; Jimenez-Cruz et al., 2010). On the contrary, some studies have found no association between consuming foods high in fats, sugars, and unprocessed and being overweight among children aged 6 to 23 months, particularly among countries with a low prevalence of childhood overweight (Budree et al., 2017; Pries, Filteau, et al., 2019). The possible reason for these differences might be different feeding practices among under-five children by country, differences in socio-economic status, cultural differences in feeding preferences, and different nutrition assessment methods (Min et al., 2021; Otitoola et al., 2021).

Malawi has programs such as the Malawi National Multi-Sector Nutrition Policy 2018–2022 that ensure that under-five-year-old children have optimal feeding (Ministry of Health, Malawi, 2018). However, the implementation and adoption of these policies tend to be challenging among the poor and rural households. The lack of access to quality healthy foods among the poor and with the rise in urbanization, children are mostly fed locally-made processed foods, which are a major source of unhealthy eating (Kalimbira & Gondwe, 2015; Casari et al., 2022). Evidence from this study support recent alarms made by public health specialist on the growing prevalence of unhealthy eating which is resulting in high childhood overweight in SSA. There is a need to intensify interventions that promote appropriate child feeding practices in Malawi and the region and to ensure the uptake of the guidelines at all socio-economic levels.

Findings in this study are subject to constraint in that there was no temporal association between unhealthy feeding and being overweight. Even though we did not have temporality, our study provides more robust findings than other studies that have used observational studies because we used covariate adjustment methods based on the propensity score. There is a possibility that other unmeasured factors would have to be controlled in terms of the propensity score matching. However, our sensitivity analysis showed that the estimated effect could not have been affected by the unmeasured factors. Another limitation was that unhealthy foods were based on self-reported information and a 24-hour recall of the type of dietary food a child was given which may result in recall bias.

Conclusions

The findings from this study revealed the adverse effect of unhealthy feeding on child overweight. Child feeding interventions in sub-Saharan African should been encouraged and supported to reduce the consumption of unhealthy foods to address the problem of child overweight in the region. .

Acknowledgements We acknowledge the DHS team for making the MDHS data available for use. Special thanks to South Africa Medical Research Council for hosting and supporting my research visits.

Author contribution HST and SOMM contributed equally to this manuscript. ?DM helped with the concept and direction of the manuscript. All authors have read and approved the final manuscript.

Funding This work was supported through the DELTAS Africa Initiative, Sub-Saharan Africa Consortium for Advanced Biostatistics Training (SSACABT). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)?s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa?s Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [SSACABT] and the UK government. The views expressed in this publication are those of the author(s) and not those of AAS, NEPAD Agency, Wellcome Trust or the UK government. This work was further supported by the L?Or?al-UNESCO for Women in Science Sub-Saharan Africa Young Talents Programme.

Data Availability The data that support the findings of this study are available upon request from the Demographic and Health Survey (DHS) website. Upon approval, full access is granted to all unrestricted survey datasets.

Declarations

Ethics approval and consent to participate The Malawi Health Research Committee determined that ethical approval was not deemed necessary in this study considering the fact that the study used data from a research study already approved by an ethical research committee. According to the 2015-16 MDHS report [11], the MDHS study was ethically approved by Malawi Health Research Committee, Institutional Review Board of ICF Macro, and Centre for Disease and Control (CDC) in Atlanta, GA, USA and Prevention IRB. Informed consent was obtained from all eligible persons.

Conflict of interest The authors declare that they have no conflict of

interest.

References

Aakvik, A. (2001). Bounding a matching estimator: the case of a Norwegian training program. *Oxford Bulletin of Economics and Statistics*, *63*(1), 115–143

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *http://i%3E/10.1080/00273171.2011.568786*.https://doi.org/10.1080/00273171.2011.568786

Austin, P. C., & Mamdani, M. M. (2006). A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Statistics in Medicine*, *25*(12), 2084–2106

Biadgilign, S., Mgutshini, T., Haile, D., Gebremichael, B., Moges, Y., & Tilahun, K. (2017). Epidemiology of obesity and overweight in sub-Saharan Africa: a protocol for a systematic review and meta-analysis. *British Medical Journal Open*, *7*(11), https://doi.org/10.1136/BMJOPEN-2017-017666

Bowker, A. H. (1948). A Test for Symmetry in Contingency Tables. *Journal of the American Statistical Association*, *43*(244), 572–574. https://doi.org/10.1080/01621459.1948.10483284

Budree, S., Goddard, E., Brittain, K., Cader, S., Myer, L., & Zar, H. J. (2017). Infant feeding practices in a South African birth cohort-A longitudinal study. *Maternal & Child Nutrition*, *13*(3), https://doi.org/10.1111/MCN.12371

Casari, S., Paola, M., Di, Banci, E., Diallo, S., Scarallo, L., Renzo, S., Gori, A., Renzi, S., Paci, M., de Mast, Q., Pecht, T., Derra, K., Kaboré, B., Tinto, H., Cavalieri, D., & Lionetti, P. (2022). Changing Dietary Habits: The Impact of Urbanization and Rising Socio-Economic Status in Families from Burkina Faso in Sub-Saharan Africa. *Nutrients 2022*, *14*(9), 1782. https://doi.org/10.3390/NU14091782. *14*

English, L. K., Obbagy, J. E., Wong, Y. P., Butte, N. F., Dewey, K. G., Fox, M. K., Greer, F. R., Krebs, N. F., Scanlon, K. S., & Stoody, E. E. (2019). Types and amounts of complementary foods and beverages consumed and growth, size, and body composition: a systematic review. *The American Journal of Clinical Nutrition*, *109*(Suppl_7), 956S–977S. https://doi.org/10.1093/ajcn/nqy281

Gu, X. S., & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics*, *2*(4), 405–420

Haider, R., & Saha, K. K. (2016). Breastfeeding and infant growth outcomes in the context of intensive peer counselling support in two communities in Bangladesh. *International Breastfeeding Journal*, *11*(1), 18. https://doi.org/10.1186/s13006-016-0077-6

Jimenez-Cruz, A., Bacardi-Gascon, M., Pichardo-Osuna, A., Mandujano-Trujillo, Z., & Castillo-Ruiz, O. (2010). Infant and toddlers' feeding practices and obesity amongst low-income families in Mexico. *Asia Pacific Journal of Clinical Nutrition*, *19*(3), 316–323

Kalimbira, A., & Gondwe, E. (2015). Consumption of sweetened beverages among school-going children in a densely populated township in Lilongwe, Malawi. *Malawi Medical Journal: The Journal of Medical Association of Malawi*, *27*(2), 55–59. https://doi.org/10.4314/mmj.v27i2.5

Kansra, A. R., Lakkunarajah, S., & Jay, M. S. (2020). Childhood and Adolescent Obesity: A Review. *Frontiers in Pediatrics*, *8*, 581461. https://doi.org/10.3389/fped.2020.581461

Lang Morović, M., & Musić Milanović, S. (2019). Breastfeeding Duration as a Predictor of Childhood Lifestyle Habits, Overweight and Obesity in Second- and Third-Grade Schoolchildren in Croatia. *Acta Clinica Croatica*, *58*(3), 481–490. https://doi.org/10.20471/acc.2019.58.03.12

Martin, R. M. (2001). Commentary: Does breastfeeding for longer cause children to be shorter? *International Journal of Epidemiology*, *30*(3), 481–484. https://doi.org/10.1093/ije/30.3.481

Min, K., Wang, J., Liao, W., Astell-Burt, T., Feng, X., Cai, S., Liu, Y., Zhang, P., Su, F., Yang, K., Sun, L., Zhang, J., Wang, L., Liu, Z., & Jiang, Y. (2021). Dietary patterns and their associations with overweight/obesity among preschool children in Dongcheng District of Beijing: a cross-sectional study. *Bmc Public Health*, *21*(1), 223. https://doi.org/10.1186/s12889-021-10240-x

Normand, S. L. T., Landrum, M. B., Guadagnoli, E., Ayanian, J. Z., Ryan, T. J., Cleary, P. D., & McNeil, B. J. (2001). Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *Journal of Clinical Epidemiology*, *54*(4), 387–398

NSO (2017). *Malawi Demographic and Health Survey 2015-16 Report*

O'Connor, T. M., Yang, S. J., & Nicklas, T. A. (2006). Beverage intake among preschool children and its effect on weight status. *Pediatrics*, *118*(4), https://doi.org/10.1542/peds.2005-2348

Otitoola, O., Oldewage-Theron, W., & Egal, A. (2021). Prevalence of overweight and obesity among selected schoolchildren and adolescents in Cofimvaba, South Africa. *South African Journal of Clinical Nutrition*, *34*(3), 97–102. https://doi.org/10.1080/16070658.2020.1733305

Pries, A. M., Filteau, S., & Ferguson, E. L. (2019). Snack food and beverage consumption and young child nutrition in low- and middle-income countries: A systematic review. *Maternal and Child Nutrition*, *15*(S4), https://doi.org/10.1111/mcn.12729

Pries, A. M., Rehman, A. M., Filteau, S., Sharma, N., Upadhyay, A., & Ferguson, E. L. (2019). Unhealthy Snack Food and Beverage Consumption Is Associated with Lower Dietary Adequacy and Length-for-Age z-Scores among 12-23-Month-Olds in Kathmandu Valley, Nepal. *The Journal of Nutrition*, *149*(10), 1843–1851. https://doi.org/10.1093/JN/NXZ140

Rosenbaum, P. R. (2005). Sensitivity Analysis in Observational Studies. *Statistics Reference Online*, *4*, 1809–1814

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Rosenbaum, P. R., & Small, D. S. (2017). An adaptive Mantel–Haenszel test for sensitivity analysis in observational studies. *Biometrics*, *73*(2), 422–430

Rousham, E. K., Goudet, S., Markey, O., Griffiths, P., Boxer, B., Carroll, C., Petherick, E. S., & Pradeilles, R. (2022). Unhealthy Food and Beverage Consumption in Children and Risk of Overweight and Obesity: A Systematic Review and Meta-analysis. *Advances in Nutrition (Bethesda, Md.)*. https://doi.org/10.1093/advances/nmac032

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice.Biometrics,249–264

Sahoo, K., Sahoo, B., Choudhury, A. K., Sofi, N. Y., Kumar, R., & Bhadoria, A. S. (2015). Childhood obesity: causes and consequences. *Journal of Family Medicine and Primary Care*, *4*(2), 187–192. https://doi.org/10.4103/2249-4863.154628

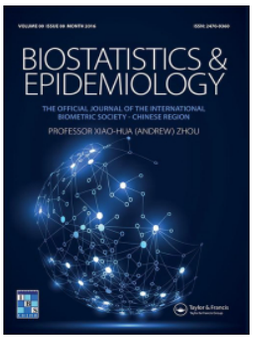
Tuyet, L. T., Thu, N. T. T., Hoai, N. T. T., Huong, N. T. L., Dung, L. T. T., & Khanh, D. N. (2019). Double Burden of Nutrition and some Eating Habits Characteristics of Preschool Children in Nam Hong Commune, Dong Anh district, Hanoi, 2018. *VNU Journal of Science: Medical and Pharmaceutical Sciences*, *35*(2), 68–77. https://doi.org/10.25073/2588-1132/VNUMPS.4175

World Health Organization. (2019). UNICEF-WHO low birthweight estimates: levels and trends 2000-2015 (No. WHO/NMH/NHD/19.21). World Health Organization.

WHO (2006). *WHO Child Growth Standards based on length/height, weight and age WHO MULTICENTRE GROWTH REFERENCE STUDY GROUP*. https://doi.org/10.1080/08035320500495548

WHO, I (2018). *Obesity and overweight*

World Health Organization (2008). *Young Child Feeding Global Strategy for Infant and*. http://apps.who.int/iris/bitstream/10665/42590/1/9241562218.pdf



Biostatistics & Epidemiology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbep20

A marginal structural model for estimation of the effect of HIV positivity awareness on risky sexual behavior

Halima S. Twabi, Samuel O. M. Manda, Dylan S. Small & Hans-Peter Kohler

To cite this article: Halima S. Twabi, Samuel O. M. Manda, Dylan S. Small & Hans-Peter Kohler (2022) A marginal structural model for estimation of the effect of HIV positivity awareness on risky sexual behavior, Biostatistics & Epidemiology, 6:2, 309-326, DOI: 10.1080/24709360.2023.2171537

To link to this article: https://doi.org/10.1080/24709360.2023.2171537

Published online: 28 Feb 2023.

BIOSTATISTICS & EPIDEMIOLOGY
2023, VOL. 6, NO. 2, e2171537 (18 pages)
https://doi.org/10.1080/24709360.2023.2171537

Taylor & Francis
Taylor & Francis Group

A marginal structural model for estimation of the effect of HIV positivity awareness on risky sexual behavior

Halima S. Twabi [a], Samuel O. M. Manda [b], Dylan S. Small [c] and Hans-Peter Kohler [d]

[a]Department of Mathematical Science, University of Malawi, Zomba, Malawi; [b]Department of Statistics, University of Pretoria, Gauteng, Pretoria, South Africa; [c]Department of Statistics, University of Pennsylvania, Philadelphia, PA, USA; [d]Department of Sociology, University of Pennsylvania, Philadelphia, PA, USA

ABSTRACT

In this paper, a Marginal Structural Model (MSM) with inverse probability of treatment weights was used to estimate the causal effect of HIV positivity awareness on condom use and multiple sexual partners using data from the Malawi Longitudinal Study of Families and Health (MLSFH). Cumulative awareness of HIV positivity was measured as the number of times an individual was aware of their positive HIV status. Awareness of HIV positivity was associated with increased condom use (OR=2.22, 95%: (1.79, 2.75)). Only among women was it associated with multiple sexual partners (OR=1.76, 95%: (1.36, 2.28)). The use of MSM (over standard regression models for repeated measures) should be encouraged as it is more suited for assessing the cumulative treatment effects while controlling for time-varying confounders in longitudinal studies. There is a need to up-scale interventions that promote HIV testing, awareness of HIV status, and prevention of HIV transmission.

ARTICLE HISTORY
Received 10 January 2022
Accepted 15 January 2023

KEYWORDS
Marginal structural models; HIV positivity; condom use; multiple sexual partners

1. Introduction

Risky sexual behaviors such as premarital sex, multiple sexual partners, and unprotected sex have been shown to be major contributors to new HIV infections and account for the high HIV transmission in sub-Saharan Africa [1–4]. With the increase in programs that promote HIV testing and HIV status awareness, research interest is now shifting to assess the effect of HIV status awareness on risky sexual behavior [5–7]. Most of the empirical evidence on the association between awareness of HIV status and risky sexual behavior comes from using cross-sectional studies [8–11], which could be limited as the exposure and outcome are observed simultaneously. Longitudinal studies provide the temporal dependence between the exposure and outcome to estimate the causal effect. Thus, there is a need for awareness of HIV status (exposure) to precede sexual behavior (outcome) to estimate the causal effect correctly.

CONTACT Halima S. Twabi htwabi@unima.ac.mw University of Malawi, P.O. Box 280, Zomba, Malawi

Gender, educational level, social-economic status, age, knowledge, and attitudes towards HIV are some factors associated with awareness of HIV status and risky sexual behaviors [1, 12–16]. In longitudinal studies, an individual's age and years of education are time-varying. Additionally, prior awareness of HIV positivity could affect school progression due to stigmatization associated with low self-confidence, anxiety, and depression [17–21]. Thus, facing a challenge of treatment-confounder feedback [22]. Such complex confounding should be taken into account in all analyses aimed at estimating the effect of awareness of HIV positivity on risky sexual behaviors.

In the presence of time-dependent confounders, standard regression models for repeated measures could result in biased treatment effect estimates resulting in erroneous conclusions [23–26]. Marginal structural models (MSM) could be used to estimate the causal effect of time-varying treatment while controlling for time-varying confounders in longitudinal studies [23]. Among the many ways of adjusting for imbalances in confounder variables when using MSM methods, the inverse probability of treatment weights (IPTW) is mostly used to estimate the treatment causal effect. The weights are obtained as a product of the inverse of the probability of treatment exposure conditional on past treatment exposures and confounders at each time point. Using HIV related data from the Malawi Longitudinal Study of Families and Health (MLSFH), this study sought to estimate the causal effect of HIV positivity awareness on condom use and multiple sexual partners.

2. Materials and methods

2.1. The Malawi longitudinal study of families and health (MLSFH)

The Malawi Longitudinal Study of Families and Health (MLSFH) which collects socio-economic, family dynamics and health information including HIV for up to 4000 couples in the 120 villages across the three districts were used [27–29]. For this paper, 2540 individuals enrolled in 2004 and followed up at eight time points namely; 2006, 2008, 2010, 2012, 2013, 2017, 2018, and 2019 were studied.

2.1.1. Exposure

This study concentrated on individuals who were aware of their HIV test results. In each year, awareness of HIV positivity was defined as an individual's knowledge of their positive HIV test result. Over the entire period from 2004 to 2019, the cumulative measure of awareness of an individual's HIV positivity was measured as the number of times the individual was aware of their positive HIV test result.

2.1.2. Outcome(s)

In this study, risky sexual behavior was measured by the number of sexual partners and condom use at last sex. Condom use at last sex was categorized as 1 (= Yes) of 0 (= No) and for number of sexual partners, it was 1 ($\geq$ 2) or 0 (if 0, 1).

2.1.3. Confounders

The following predictor variables were considered as baseline confounders: region, religion, wealth, given an incentive to collect HIV test results, distance to HIV testing and counseling (HTC) centers, perceived risk of HIV infection, gender, and comprehensive

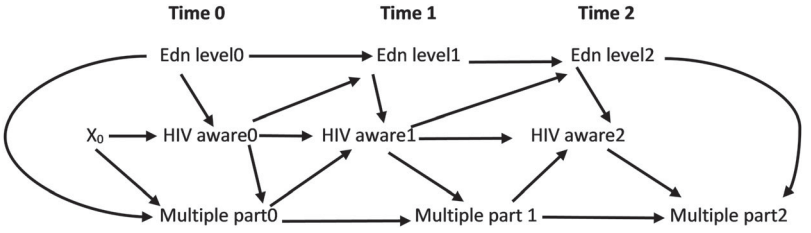


Figure 1. Directed Acyclic Graph (DAG) on HIV status awareness on multiple sexual partners for three time periods where X_0 denotes the baseline confounders.

HIV knowledge. Age in years and education level (0 to 12 years) were considered as time-varying confounders. Figure 1 illustrates this relationship.

2.2. Marginal structural models

Suppose for individual i, $i = 1, \ldots, n$, a binary exposure (awareness of HIV status) is observed, $A_t = (0 -$ awareness of HIV negativity, $1 -$ awareness of HIV positivity), at time t for $t = 1(2004), \ldots, 9(2019)$. A_1 represents awareness of HIV status at baseline ($t = 1$). For individual i, $\mathbf{X}_t$ represents a vector of time-varying confounders that may be associated with A_t, where $\mathbf{X}_1$ are the baseline confounders. Further, $\bar{A}_t = (A_1, A_2, \ldots, A_9)$, and $\bar{\mathbf{X}}_t = (X_1, X_2 \ldots, X_9)$ denote the observed histories of treatment, and observed confounders from baseline through the follow-up times (2006, 2008, …, 2019), respectively. The vector $\bar{a}_t = (a_1, a_2, \ldots, a_9)$ denotes the history of the possible values of $\bar{A}_t$ at each time point. Let $Y_t(\bar{a}_t)$ be the potential condom use (or multiple sexual partners) at time t that would have been observed had an individual been aware of their HIV status history (either negative or positive). Then, the estimate of interest is the difference in the means of the potential outcomes $E[Y_t(\bar{a}_t)]$ between different awareness of HIV positivity exposures, $\bar{a}_t$, at time t. For the two binary outcomes, $E[Y_t(\bar{a}_t)]$ is simply $P(Y_t(\bar{a}_t) = 1)$, the conditional probability of the respective outcome=1 at time t. The MSM can be modeled as:

$$E[Y_{t+1}(\bar{a}_t)] = \text{logit}(\text{P}(Y_t(\bar{a}_t) = 1)) = g(\bar{a}_t; \boldsymbol{\beta}) \tag{1}$$

where $g(\bar{a}_t; \boldsymbol{\beta})$ is a specified function of the exposure which depends on the question of interest. One would be interested in assessing the effect of being always aware of HIV positivity (i.e. $\bar{a}_t = (1, 1, \ldots, 1)$) on risky sexual behavior against never aware (i.e. $\bar{a}_t = (0, 0, \ldots, 0)$), or aware of HIV positivity at $t = 4$ and $t = 7$, or allow awareness at different times to have different treatment effects $\sum_{t=1}^{T} \beta_t a_t$. Additionally, the additive effect of awareness of HIV positivity on condom use (or multiple sexual partners) may be of interest and the cumulative positivity would be used. We modeled the MSM based on the observed awareness of HIV status and outcome at time (t) for individual i as follows;

$$\text{logit}(\text{P}(Y_{t+1} = 1 | \bar{A}_t = \bar{a}_t)) = \beta_0 + \beta_1 \sum_{k=1}^{9} A_t \tag{2}$$

Where β_0 is the unknown intercept, β_1 is the estimated effect of cumulative awareness of HIV positivity ($\sum_{k=1}^{9} A_t$) compared to HIV negativity awareness on reducing multiple

sexual partners or increase condom use over time, as a hypothesis. Equation (2) assumes that each additional period of awareness of HIV status has an additive effect (the effect over time) on change in condom use and multiple sexual partners

The parameter β_1 is estimated under the assumptions of no unmeasured confounding (exchangeability) and positivity, known as the strong ignorability assumption. Exchangeability assumes that the potential outcomes (the hypothetical outcomes under given exposure levels) are independent of the actual exposure level. For example, for the relationship between condom use and awareness of HIV status, the potential outcomes are condom use were an individual to be aware of HIV negativity, and condom use were that same individual to be aware of HIV positivity at time t (two of these potential outcomes will be counterfactual). Exchangeability would mean that these potential outcomes are independent of an individual's actual HIV awareness status after accounting for all the measured confounders, including HIV status awareness history. The positivity assumption implies that individuals have a non-zero probability of being either aware of HIV negativity or aware of HIV positivity, for all the identified covariates and past exposure history. In this study, the exposure was assumed to be time-varying and once the individual was aware of their HIV positivity, it remained constant.

2.2.1. Inverse probability treatment weights (IPTW)

The inverse probability of treatment weights (IPTWs) were used to adjust for confounding. The IPTWs were the inverse of the conditional probability of awareness of HIV positivity at time t, given prior awareness of HIV positive status and covariate history. They were calculated for each individual by first estimating the probability, $P(\mathrm{A}_t = 1|\bar{A}_{t-1}, \bar{\mathbf{X}})$, using a multivariable logistic regression model and then obtaining a product of their inverses to get the overall weight;

$$W_i^a(t) = \prod_{k=1}^{t} \frac{1}{P(A_k = 1 \mid \bar{A}_{k-1} = \bar{a}_{k-1}, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k, \mathbf{X}_1)} \tag{3}$$

where $k = 1, \ldots, t$, $\bar{A}_{k-1}$ is the past exposure history. Individuals aware of their HIV positivity were given a weight of $1/P(\mathrm{A}_t = 1|\bar{A}_{t-1}, \bar{\mathbf{X}})$ and those aware of HIV negativity were given a weight of $1/(1 - P(\mathrm{A}_t = 1|\bar{A}_{t-1}, \bar{\mathbf{X}}))$. Individuals with a small probability of awareness of HIV positivity may obtain a large weight and vice versa resulting in extreme weights. Extreme weights affect the estimated effects by inflating the variance and confidence intervals [30, 31]. Thus, extreme weights can be handled through either weight stabilization or weight truncation (trimming). In this study, the stabilized inverse probability treatment weights (sIPTW) were used. Stabilizing the weights was achieved by replacing the numerator in (3) (which is 1 in the unstabilized weights) with the probability of awareness of HIV positivity given past awareness of HIV positivity and all the baseline confounders (i.e. without time-dependent covariates). The sIPTW were calculated as follows;

$$SW_i^a(t) = \prod_{k=1}^{t} \frac{P(A_k = 1 \mid \bar{A}_{k-1} = \bar{a}_{k-1}, \mathbf{X}_1)}{P(A_k = 1 \mid \bar{A}_{k-1} = \bar{a}_{k-1}, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k, \mathbf{X}_1)} \tag{4}$$

$\bar{\mathbf{X}}_k$ is the vector of time-dependent covariate history. The numerator for the sIPTW (Equation (4)) was also estimated using a multivariable logistic regression model.

The weights described in (3) and (4) only control for time-varying confounders and exposures and are not designed to account for loss-to-follow-up (LTFU). Thus, to control for loss-to-follow-up, one could consider defining censoring weights. Individuals who refused to participate, were dead, moved out or not eligible to participate in the MLSFH study were considered as censored while individuals who were temporarily absent were uncensored. Let D_t indicate loss-to-follow-up throughout the study, then $D_t = (1 -$ censored and, $0 -$ not censored). We denote $\bar{D}_t = (D_1, D_2, \ldots, D_9)$ as the LTFU history with the assumption of no censoring at baseline ($D(1) = 0$). The weights for censoring were obtained by estimating the probability of remaining in the study (uncensored) at time t conditional on being uncensored in the past ($t-1$) and the observed confounders. The weights for LTFU are defined as;

$$SW_i^d(t) = \prod_{k=1}^{t} \frac{P(D_k = 0 \mid \bar{D}_{k-1} = 0, \bar{A}_k = \bar{a}_k)}{P(D_k = 0 \mid \bar{D}_{k-1} = 0, \bar{A}_k = \bar{a}_k, \bar{\mathbf{X}}_k = \bar{\mathbf{x}}_k)} \tag{5}$$

The final weights for individual i were calculated as a product of the stabilized and the censoring weights at time t.

$$SW_i = \mathrm{SW}_i^a(t) \times \mathrm{SW}_i^d(t) \tag{6}$$

The weight in (6) correctly accounts for confounding bias due to confounders (both time-independent and time-varying covariates) under the assumptions of strong ignorability (exchangeability and positivity), consistency and no mis-specification of the model used to estimate the treatment probabilities [32, 33].

3. Description of the data and results

3.1. Awareness of HIV status by year

The distribution of awareness of HIV status by year of collection is shown in Table 1. A majority (92.8%) of the individuals were HIV negative while 183 (7.2%) were HIV positive. In 2004, more than half of those who were HIV positive (62.8%) were aware of their HIV positivity while 68 (37.2%) were unaware of their HIV positive status. Awareness of HIV positivity increased over time among those who were HIV positive, with a majority of individuals in 2006 (96.8%) and 2008 (97.4%) being aware of their HIV positivity. All the individuals who were HIV positive in 2010 and 2012 were aware of their HIV positive status. In 2017 and 2018, 73 individuals who were HIV positive were aware of their HIV positivity status. In subsequent years from 2004, awareness of HIV positivity increased.

3.2. Assessing balance of confounders

The distribution of the baseline characteristics for the study participants across awareness of HIV positivity or negativity together with their respective associations are presented in Table 2. A number of the confounders were found to have different distributions between the HIV positivity and negativity awareness groups. Awareness of HIV positivity was also associated with having multiple sexual partners. This provided empirical support for the causal relationship hypothesized in Figure 1.

Table 1. Distribution of HIV status awareness by wave.

	Awareness of HIV status						
	HIV positivity (N=183 (7.2%))			HIV negativity (N=2357 (92.8%))			
Year	Yes	No	Subtotal	Yes	No	Subtotal	Total
2004	115 (62.8%)	68 (37.2%)	183	1626 (69.0%)	731 (31.0%)	2357	2540
2006	91 (96.8%)	3 (3.2%)	94	1568 (98.7%)	20 (1.3%)	1588	1682
2008	74 (97.4%)	2 (2.6%)	76	1448 (99.6%)	6 (0.4%)	1454	1530
2010	74 (97.4%)	2 (2.6%)	76	1418 (99.6%)	6 (0.4%)	1424	1500
2012	42 (100%)	0 (0.0%)	42	607 (97.6%)	15 (2.4%)	612	664
2013	42 (100%)	0 (0.0%)	42	606 (97.6%)	15 (2.4%)	621	663
2017	73 (98.6%)	1 (1.4%)	74	819 (100%)	0 (0.0%)	819	893
2018	73 (98.6%)	1 (1.4%)	74	819 (100%)	0 (0.0%)	819	893
2019	58 (100%)	0 (0.0%)	58	671 (100%)	0 (0.0%)	671	729

Table 2. Distribution of baseline characteristics by HIV status awareness of study participants in the MLSFH for year 2004.

	Awareness of HIV status		
	HIV positivity	HIV negativity	p-value
Total (*n*=1741)	115 (6.6%)	1626 (93.4%)	
	Gender		
Women	67 (58.3)	870 (53.5%)	
Men	48 (41.7%)	756 (46.5%)	0.323
	Age category		
15–44 years	83 (72.2%)	1182 (72.7%)	
45 years above	32 (27.8%)	444 (27.3%)	0.525
	Wealth		
Poor	45 (45%)	547 (41.9%)	
Medium/rich	55 (55.0%)	759 (58.1%)	0.543
	Education		
None	23 (23.5%)	355 (24.8%)	
Some education	75 (76.5%)	1076 (75.2%)	0.766
	Religion		
Christian	53 (48.6%)	859 (55.4%)	
Muslim or other	56 (51.4%)	677 (44.1%)	0.043
	Incentive		
No	34 (45.9%)	482 (48.5%)	
Yes	40 (54.1%)	512 (51.5%)	0.044
	Region		
North	43 (37.4%)	536 (32.9%)	
Central	51 (44.4%)	600 (36.9%)	
South	21 (18.3%)	490 (30.2%)	0.025
	Condom Use		
No	68 (73.9%)	1001 (74.5%)	
Yes	24 (26.1%)	342 (25.5%)	0.895
	Multiple partners		
No	20 (19.2%)	535 (39.1%)	
Yes	84 (80.8%)	832 (60.9%)	< 0.001

The study had individual dropouts either due to death, migration, or not being eligible to participate. Thus, the distribution of the purported confounders differed by loss-to-follow-up (LTFU) for the succeeding year of 2004 was assessed, and the findings are presented in Table A1 in the Appendix. The findings show differences in distribution between the LTFU levels among some of the measured confounders; hence, we controlled for loss-to-follow-up in the analysis.

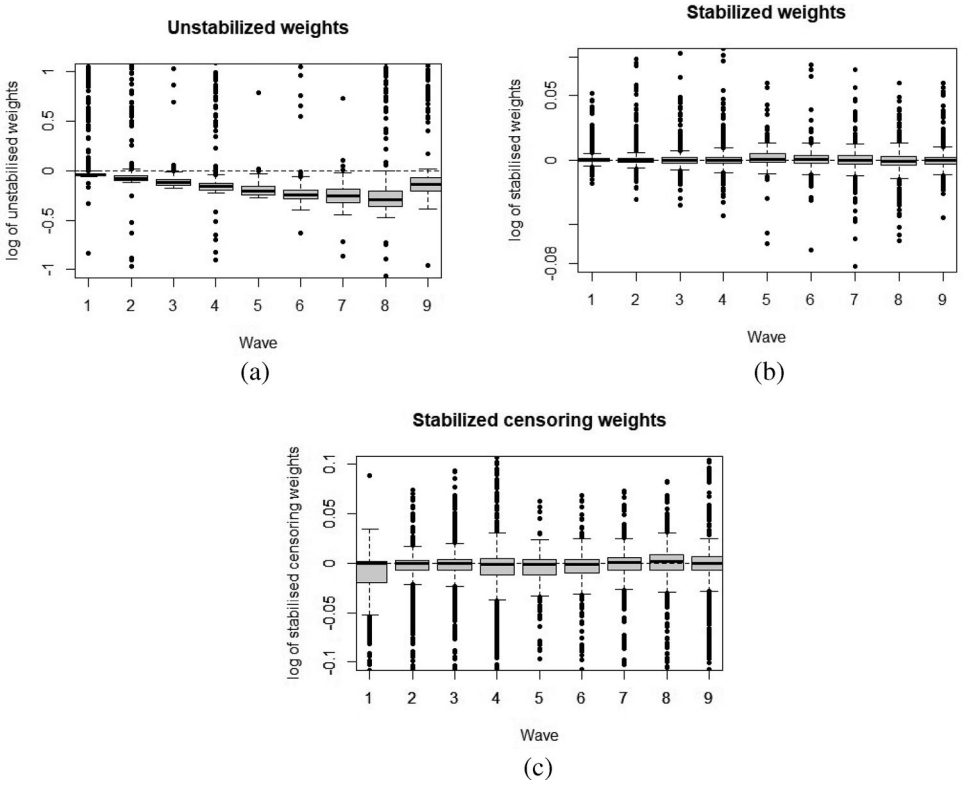


Figure 2. Distribution of the log of non-stabilized, censoring, and stabilized weights.

Table 3. Distribution of inverse probability weights.

Weights	Mean (SD)	Median	Interquartile range	Min	Max
Unstabilized	0.89 (0.5)	0.88	(0.81, 0.95)	0.0	16.40
Stabilized	1.0 (0.01)	1.0	(0.99, 1.00)	0.8	1.2
Censoring	0.99 (0.18)	0.99	(0.95, 1.03)	0.16	5.97

3.3. Distribution of weights

Table 3 presents the distribution of the unstabilized, stabilized and censoring weights. The unstabilized weights had a mean of 0.89, a median of 0.88, a standard deviation of 0.5, and an inter-quartile range of (0.81, 0.95). The maximum weight was 16.4, implying that an individual with the largest weight contributed about 16 people to the pseudo-population. The unstabilized weights had a huge variability across the time points, see Figure 2(a). Conversely, the stabilized weights had a median and mean centered at 1.0 across all time points, as shown in Figure 2(b). The range narrowed to a minimum of 0.99 and a maximum of 1.00. The maximum weight was 1.2, which corresponded to a contribution of 1 person to the pseudo-population. The censoring weights had a mean of 0.99 and median of 0.99 and ranged between 0.95 and 1.03, see Figure 2(c). The stabilized weights were less variable (SD=0.01) and the mean and median were centered around 1.0 as required, compared to the censoring and the unstabilized weights.

The distributions of the weights used for the MSM are shown in Figure 2a (Distribution of the log of unstabilized and stabilized weights) for several follow-up times. The weights are presented as a logarithmic form for display purposes. The unstabilized weights produced a skewed distribution with extreme weights as compared to the stabilized weights.

3.4. Effect of HIV status awareness on risky sexual behavior

A weighted (IPTW) marginal logistic regression model was used to estimate the awareness of HIV positivity effects. The weights were calculated using the `IPW` package [34]. For modeling repeated multiple partners and condom use measurements, independent and AR(1) working correlation matrices were assumed and we estimated the variance using 'robust' standard errors using the `geepack` package in R [35]. An independent working correlation matrix was used in our analyses because other correlation structures defined using the generalized estimating equations (GEEs) have been shown to produce biased treatment effect estimates when past outcomes affect future exposures (in our example having multiple sexual partners may influence an individuals decision to know their HIV positivity) [36, 37]. The weighted models were compared to the estimated treatment effects obtained from the standard multivariable binary logistic model were we also adjusted for the effect of year.

3.4.1. Standard multivariable logistic regression model

Table 4 presents the odds ratio for the association between awareness of HIV positivity and condom use or multiple sexual partners. Condom use was significantly associated with awareness of HIV positivity (OR: 2.30; 95% CI=(1.82, 2.90)), (Table 4(a)). Additionally, men (OR: 1.72; 95% CI=(1.50, 1.97)); individuals with a primary (OR: 1.32; 95% CI=(1.08, 1.62)) and secondary education (OR: 1.73; 95% CI=(1.31, 2.29)); individuals from medium or rich households (OR: 1.16; 95% CI=(1.02, 1.32)) and those from the southern region (OR: 2.12; 95% CI=(1.34, 3.37)) were more likely to use condoms.

Table 4(b) presents the association between multiple sexual partners and awareness of HIV positivity adjusting for the measured confounders. Individuals aware of HIV positivity were more likely to have multiple partners (OR: 1.52; 95% CI=(1.17, 1.98)). Men were more likely to have multiple partners than women (OR: 3.11; 95% CI=(2.67, 3.62)). On the other hand, individuals from medium or rich households (OR: 0.81; 95% CI=(0.71, 0.93)) were less likely to have multiple sexual partners. In addition, over time individuals were less likely to have multiple partners (OR: 0.71; 95% CI=(0.68, 0.75)).

3.4.2. Marginal structural model

Both the unstabilized and stabilized weights were used to fit the MSM model. Only the results for the stabilized weights are reported (see, Table 5) as the distribution of the weights were more stable (the results for the unstabilized weights are reported in Table A2 in the Appendix). In using the MSM, the higher the number of times a person was aware of their HIV positivity resulted in an increase in condom use (OR: 2.22; 95% CI=(1.79, 2.75)) compared to those individuals aware of HIV negativity (see Table 5(a)). However, the cumulative effect of awareness of HIV positivity on multiple sexual partners was not significant (OR: 1.06; 95% CI=(0.83, 1.35)) (see Table 5(b)).

Table 4. Effect estimates (Odds ratios, (95% confidence Interval)) using standard multivariable logistic model.

Covariates	Odds ratio	95% Conf. Int.
(a) Condom Use		
Intercept	1.62	(1.23, 2.14)
HIV aware (HIV negativity vs HIV positivity)	2.30	(1.82, 2.90)
Gender (women vs men)	1.72	(1.50, 1.97)
Age (in years)	0.43	(0.37, 0.50)
Wealth (poor vs medium/rich)	1.16	(1.02, 1.32)
Education (none vs primary)	1.32	(1.08, 1.62)
Education (none vs secondary/Tertiary)	1.73	(1.31, 2.29)
Religion (Muslim or other vs Christian)	1.02	(0.88, 1.18)
Region (North vs Central)	0.99	(0.74, 1.32)
Region (North vs South)	2.12	(1.34, 3.37)
Year	0.99	(0.96, 1.02)
(b) Multiple Sexual Partners		
Intercept	0.31	(0.23, 0.42)
HIV aware (HIV-negativity vs HIV positivity)	1.52	(1.17, 1.98)
Gender (women vs men)	3.11	(2.67, 3.62)
Age (in years)	1.00	(0.99, 1.01)
Wealth (poor vs medium/rich)	0.81	(0.71, 0.93)
Education (none vs primary)	1.07	(0.89, 1.30)
Education (none vs secondary/Tertiary)	0.88	(0.62, 1.25)
Religion (Muslim or other vs Christian)	1.16	(0.98, 1.37)
Region (North vs Central	1.22	(0.90, 1.68)
Region (North vs South)	0.95	(0.56, 1.63)
Year	0.71	(0.68, 0.75)

Note: 95% Conf. Int. – 95% Confidence Interval.

Table 5. Effect Estimates of awareness of HIV positivity on risky sexual behavior using stabilized weights.

	Odds ratio	95% CI	*p*-value
(a) Condom Use			
Intercept	0.11	(0.10, 0.14)	< 0.001
HIV negativity	1.00	–	–
HIV positivity	2.22	(1.79, 2.75)	< 0.001
(b) Multiple Sexual Partners			
Intercept	0.24	(0.21, 0.27)	< 0.001
HIV negativity	1.00	–	–
HIV positivity	0.96	(0.83, 1.35)	0.63

Notes: (i) We controlled for the following confounders:- gender, wealth, age, education, time (wave), religion, comprehensive HIV knowledge, distance to VCT, received an incentive, perceived partners HIV status, likelihood has HIV, region.
(ii) 95% CI – 95% Confidence Interval.

Based on the results, the standard logistic regression model showed an increased effect on multiple sexual partners. On the other hand, the MSM produced a decreased effect of the multiple sexual partner estimates. From the findings, the unstabilised weighted model produced treatment effect estimates with wider confidence intervals implying large standard errors due to the high variability in the weights, unlike the estimated treatment effects from the stabilized weighted model.

Table 6. Effect estimates of awareness of HIV positivity by gender.

	Men		Women	
	Odds ratio	95% CI	Odds ratio	95% CI
		(a) Condom Use		
Intercept	0.19	(0.14, 0.26)	0.10	(0.07, 0.11)
		Awareness		
HIV negativity	1.00	–	1.00	–
HIV positivity	2.32	(1.66, 3.24)	2.24	(1.70, 2.95)
		Multiple partners		
No	1.00	–	1.00	–
Yes	1.45	(1.19, 1.76)	1.29	(1.02, 1.63)
		(b) Multiple Sexual Partners		
Intercept	0.60	(0.46, 0.77)	0.14	(0.12, 0.17)
		Awareness		
HIV negativity	1.00	–	1.00	–
HIV positivity	0.76	(0.49, 1.18)	1.76	(1.36, 2.28)
		Condom use		
No	1.00	–	1.00	–
Yes	1.41	(1.18, 1.67)	1.28	(1.01, 1.62)

Note: 95% CI – 95% Confidence Interval.

3.4.3. The effect of awareness of HIV positivity status by gender

The effect of awareness of HIV positivity on condom use and multiple sexual partners was further stratified by gender (Table 6). Both men (OR= 2.32; 95% CI:(1.66, 3.24)) and women (OR= 2.24; 95% CI:(1.70, 2.95)) with a higher number of times aware of their HIV positivity were more likely to use condoms (Table 6(a)). However, only the women aware of their HIV positivity were more likely to have multiple sexual partners (OR= 1.76; 95% CI:(1.36, 2.28)) (Table 6(b)).

Further assessing the effect of condom use on multiple sexual partners and vice versa stratified by gender showed that men (OR= 1.45; 95% CI:(1.19, 1.76)) and women (OR= 1.29; 95% CI:(1.02, 1.63)) who had multiple sexual partners were more likely to use condoms (Table 6(a)). Similarly, men (OR= 1.41; 95% CI:(1.18, 1.67)) and women (OR= 1.28; 95% CI:(1.01, 1.62)) who used condoms were more likely to have multiple partners (Table 6(b)). This may imply the possible correlation between the risky sexual behavior outcomes.

3.5. Sensitivity analysis

The MSM using IPTW was employed under the assumption of no unmeasured confounding and this cannot be tested. Some unobserved confounders such as disclosure of partner's HIV status, alcohol use, and pre-and-post counseling may confound the estimated effects of awareness of HIV positivity on condom use and multiple sexual partners [38]. In this study, the impact of unobserved confounders on the estimated effects were assessed using the so-called ('*evidence of causality*') or *E*-value statistic [39, 40]. It was found that the effect on multiple sexual partners could have been affected greatly with a smaller impact of unmeasured confounders, but a larger impact could have been required to alter the effect on condom use (Table A3).

4. Discussion

Marginal structural models can be used to estimate the effect of an exposure on an outcome by adjusting for time-varying confounding using an inverse probability of treatment weights (IPTW) in longitudinal studies. Using this technique, the study set out to estimate the causal effect of awareness of HIV positivity on condom use and multiple sexual partners using data from the Malawi Longitudinal Study of Families and Health (MLSFH). Awareness of HIV positivity was associated with increased condom use. It was significantly associated with multiple sexual partners among women only.

The finding of an association between awareness of HIV positivity and increased condom use has been found in some previous studies [8, 16, 41–43]. This could indicate the effectiveness of policies and interventions in Malawi, such as the Malawi National Condom Strategy 2015-2020, the National HIV prevention strategy, and the National Sexual Reproductive Health and Rights (SRHR) policy [44], that are aimed at promoting safer sex to prevent HIV transmissions. Additionally, an increase in condom use among individuals aware of HIV positivity could also be attributed to individuals having multiple sexual partners and opting to use condoms for protection [45, 46]. Thus, future studies may consider accounting for the dependency between condom use and multiple sexual partners when assessing the effect of HIV positivity awareness on risky sexual behavior.

We found that among women, awareness of HIV positivity was associated with multiple sexual partners. It has been reported elsewhere that due to high poverty levels in the region, most women have two or more sexual partners after knowing their positive HIV status [47–49]. While this may partly explain our findings, the MLSFH is a rural based study, and individuals often self-report sexual behaviors in a culturally appropriate manner for approval (social desirability) [50, 51]. Thus, individuals may over-report desirable sexual behaviors (having few partners) and under-report undesirable sexual behaviors (having more partners). This biased self-reporting could have affected the responses on the number of sexual partners. Further studies are needed to ascertain differentials in sexual behavior reporting between men and women.

The study found that both the standard binary logistic regression and the MSM showed an increase in condom use among individuals aware of their HIV positivity. Some studies have also shown similar effects between the standard regression model and the MSM [52–55]. This may indicate less influence of prior exposure and time-varying confounders on the association between HIV positivity awareness and condom use. In our study, this could be attributed to education in years not being an important time-varying confounder for the association. However, using the standard multivariable logistic model, awareness of HIV positivity was associated with increased multiple sexual partners, a finding only found among women when using the MSM approach. We believe the differences in the finding could be due to the use of the inverse probability treatment weights and correct adjustment for time-dependent confounders (age and education level) [23, 25, 56].

The study's strength is that a statistical approach that controls for time-varying confounders was used. In addition, the study comprised rich data which had information on individuals followed up for over a decade, providing a change in trends and behaviors among the rural community in Malawi. The study had several limitations, including the high attrition rate within the data, which may have resulted in a loss of information. Repeated HIV testing in the subsequent years may have resulted in more individuals being

aware of their HIV status, which may have affected changes in sexual behaviors resulting in individuals using condoms or reducing the number of sexual partners over the years. However, using the inverse probability treatment weights, past HIV status awareness was controlled to ensure the estimated effects were not confounded by past awareness. Also, data on condom use and multiple partners was self-reported. As alluded to above, these data could be subject to biases from over-and under-reporting [50, 57–60] and may also affect the temporal relationship between the HIV status awareness and the outcomes due to recalling information from the past years.

The MSM used was considered the most appropriate model fit for the data. However, model misspecification was possible since we only controlled for purported confounders available within the data. Some important predictors, such as knowing the partner's HIV status for condom use and alcohol use for multiple sexual partners, were not collected during earlier years. Thus, controlling for purported confounders may have been insufficient. The findings from the sensitivity analysis showed that the effect on multiple sexual partners could be biased due to unmeasured confounders, unlike the effect on condom use. Thus, the findings on the causal effect of awareness of HIV positivity on multiple sexual partners need to be interpreted with caution. Regardless of these limitations, this study has demonstrated the capabilities of marginal structural models to estimate treatment effects in longitudinal studies as it is more suited for controlling measured confounders. Apart from using the IPTWs to control for time-varying confounders and exposures when using the MSM to estimate treatment effects, an alternative confounder balance approach could be to use an instrumental variable(s) to control for both measured and unmeasured confounders [61, 62].

5. Conclusion

The marginal structural model with inverse probability of treatment weight has the capability to correctly estimate treatment effects for longitudinal data with time-varying treatments and confounders. Using HIV data from the Malawi Longitudinal Study of Families and Health (MLSFH), we estimated the effect of awareness of HIV positivity on condom use and number of sexual partners. The study found that awareness of HIV positivity was associated with increased condom use, which supports the need for sustained interventions that promote HIV testing and awareness of HIV status.

Acknowledgments

This work was partially supported through the DELTAS Africa Initiative, Sub-Saharan Africa Consortium for Advanced Biostatistics Training (SSACABT) (The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [SSACABT] and the UK government), and by the L'Oréal-UNESCO for Women in Science Sub-Saharan Africa Young Talents Programme through the Young Talents award provided to Ms. Twabi. Ms. Twabi also thanks the South African Medical Research Council (SAMRC) for funding part of her research visits. The views expressed in this publication are those of the author(s) and not those of the funders (AAS, NEPAD Agency, Wellcome Trust or the UK government, L'Oreal, SAMRC). We further thank the MLSFH team led by Professor Hans-Peter Kohler for granting us access to the MLSFH data.

Data availability statement

The data is available upon request from the Principal investigator of the MLSFH that can be done through email: hpkohler@pop.upenn.edu.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

HT thanks the Biostatistics Research Unit of the South Africa Medical Research Council (SAMRC) for hosting her Ph.D research visits.

Notes on contributors

Halima S. Twabi performed data analysis, interpretation, and drafting of the manuscript.

Samuel O.M. Manda conceptualized, reviewed the statistical analysis and provided critical revisions.

Dylan S. Small helped with the concept, direction and revisions.

Hans-Peter Kohler helped with providing the substantive concept of the paper and the revisions of the manuscript. All authors have read and approved the final manuscript.

Ethical approval and consent to participate

The data collections of the MLSFH have been approved by the IRB Board at the University of Pennsylvania (IRB Protocols no. 815 016 and no. 826828), and in Malawi, the MLSFH research has been approved by the Ethics Committee of the College of Medicine, Malawi (COMREC, Protocols no. P01/12/1165 and no. P.04/17/2160) and the National Health Sciences Research Committee (NHSRC, Protocol no. 19/01/2214).

ORCID

Halima S. Twabi http://orcid.org/0000-0002-0166-9393
Samuel O. M. Manda http://orcid.org/0000-0002-9672-3312
Dylan S. Small http://orcid.org/0000-0003-4928-2646
Hans-Peter Kohler http://orcid.org/0000-0003-0639-4349

References

[1] Machira K, Maonga B, Chirwa T. Determinants of risky sexual behavioral practices among teen-girls in malawi. Open Public Health J. 2020;13(1):839–849.
[2] Unaids. Understanding fast-track: accelerating action to end the aids epidemic by 2030. UNAIDS Geneva; 2015.
[3] Wondemagegn F, Berkessa T. High level risky sexual behavior among persons living with hiv in the urban setting of the highest hiv prevalent areas in ethiopia: implications for interventions. PLoS ONE. 2020;15(11):Article ID e0242701.
[4] Ortblad KF, Musoke DK, Ngabirano T, et al. Is knowledge of hiv status associated with sexual behaviours? a fixed effects analysis of a female sex worker cohort in urban uganda. J Int AIDS Soc. 2019;22(7):Article ID e25336.
[5] Palk L, Blower S. Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the hiv epidemic in malawi. BMC Med. 2018;16(1):1–9.

[6] Adaralegbe A, Adeoti O, Ogbonda S, et al. Sexual behaviors and hiv status: a population-based study among youths and adults in tanzania. J Infect Dis Epidemiol. 2021;7:224.
[7] Rosenberg MS, Gómez-Olivé FX, Rohr JK, et al. Sexual behaviors and hiv status: a population-based study among older adults in rural south africa. J Acquir Immune Defic Syndr (1999). 2017;74(1):e9–e17.
[8] Bunnell R, Opio A, Musinguzi J, et al. Hiv transmission risk behavior among hiv-infected adults in uganda: results of a nationally representative survey. Aids. 2008;22(5):617–624.
[9] Guiella G, Madise NJ. Hiv/aids and sexual-risk behaviors among adolescents: factors influencing the use of condoms in burkina faso. Afr J Reprod Health. 2007;11(3):182–196.
[10] Huerga H, Venables E, Ben-Farhat J, et al. Higher risk sexual behaviour is associated with unawareness of hiv-positivity and lack of viral suppression–implications for treatment as prevention. Sci Rep. 2017;7(1):1–7.
[11] Ventura-Filipe EM, Newman SP. Influence of hiv positive status on sexual behavior among males. Rev Saude Publica. 1998;32:6503–513.
[12] Zuilkowski SS, Jukes MC. The impact of education on sexual behavior in sub-saharan africa: a review of the evidence. AIDS Care. 2012;24(5):562–576.
[13] Ali MM, Merdad L, Bellizzi S. Socioeconomic variations in risky sexual behavior among adolescents in 14 sub-saharan africa countries who report ever having had sex. Int J Equity Health. 2021;20(1):1–7.
[14] Kugbey N, Ayanore MA, Amu H, et al. International note: analysis of risk and protective factors for risky sexual behaviours among school-aged adolescents. J Adolesc. 2018;68:66–69.
[15] Nabukenya AM, Matovu JK. Correlates of hiv status awareness among older adults in uganda: results from a nationally representative survey. BMC Public Health. 2018;18(1):1–8.
[16] Nabukenya AM, Nambuusi A, Matovu JK. Risk factors for hiv infection among married couples in rakai, uganda: a cross-sectional study. BMC Infect Dis. 2020;20(1):1–8.
[17] Mathew RN, Manjula M. Self-esteem and self-efficacy among hiv-positive adolescents: an intervention study. HIV/AIDS Rev Int J HIV-Rel Prob. 2022;21(1):37–42.
[18] Zinyemba T, Pavlova M, Groot W. Effects of hiv on gender gaps in school attendance of children in zimbabwe: a non-linear multivariate decomposition analysis. Educ Econ. 2021;29(5):471–489.
[19] Parchure R, Jori V, Kulkarni S, et al. Educational outcomes of family-based hiv-infected and affected children from maharashtra, india. Vulnerable Child Youth Stud. 2016;11(4):332–338.
[20] Akinyemi J, Fotso A, Banjo O. Hiv and adolescents' educational attainment in south africa: disentangling the effect of infection in children and household members. S Afr J Child Health. 2018;2018(1):s4–s9.
[21] Toska E, Cluver L, Orkin M, et al. Screening and supporting through schools: educational experiences and needs of adolescents living with hiv in a south african cohort. BMC Public Health. 2019;19(1):1–10.
[22] Karim ME, Tremlett H, Zhu F, et al. Dealing with treatment-confounder feedback and sparse follow-up in longitudinal studies: application of a marginal structural model in a multiple sclerosis cohort. Am J Epidemiol. 2021;190(5):908–917.
[23] Hernan MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on cd4 count with a marginal structural model for repeated measures. Stat Med. 2002;21(12):1689–1709.
[24] Robins JM, Hernan MA. Estimation of the causal effects of time varying exposures. In: Longitudinal data analysis. Chapman and Hall/CRC; 2009. p. 553–99.
[25] VanderWeele TJ, Jackson JW, Li S. Causal inference and longitudinal data: a case study of religion and mental health. Soc Psychiatry Psychiatr Epidemiol. 2016;51(11):1457–1466.
[26] Robins JM. Causal inference from complex longitudinal data. In: Latent variable modeling and applications to causality. Springer; 1997. p. 69–117.
[27] Kohler HP, Watkins SC, Behrman JR, et al. Cohort profile: the malawi longitudinal study of families and health (mlsfh). Int J Epidemiol. 2014;44(2):394–404.
[28] Baranov V, Kohler HP. The impact of aids treatment on savings and human capital investment in malawi. Am Econ J: Appl Econ. 2018;10(1):266–306.

[29] Kohler Iliana V, Bandawe C, Ciancio A, et al. Cohort profile: the mature adults cohort of the malawi longitudinal study of families and health (mlsfh-mac). BMJ Open. 2020;10(10):Article ID e038232.
[30] Chesnaye NC, Stel VS, Tripepi G, et al. An introduction to inverse probability of treatment weighting in observational research. Clin Kidney J. 2022;15(1):14–20.
[31] Cole SR, Hernán MA, Robins JM, et al. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. Am J Epidemiol. 2003;158(7):687–694.
[32] Robins JM. Association, causation, and marginal structural models. Synthese. 1999;121(1-2):151–179.
[33] Robins JM. Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. In: Statistical models in epidemiology, the environment, and clinical trials. Springer; 2000. p. 95–133.
[34] van der Wal WM, Geskus RB, et al. ipw: an r package for inverse probability weighting. J Stat Softw. 2011;43(13):1–23.
[35] Højsgaard S, Halekoh U, Yan J. The r package geepack for generalized estimating equations. J Stat Softw. 2005;15(1):1–11.
[36] Keogh RH, Daniel RM, Vanderweele TJ, et al. Analysis of longitudinal studies with repeated outcome measures: adjusting for time-dependent confounding using conventional methods. Am J Epidemiol. 2017;187(5):1085–1092.
[37] Sullivan Pepe M, Anderson GL. A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. Comm Statist Simulation Comput. 1994;23(4):939–951.
[38] Crepaz N, Hart TA, Marks G. Highly active antiretroviral therapy and sexual risk behavior: a meta-analytic review. Jama. 2004;292(2):224–236.
[39] VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. Ann Intern Med. 2017;167(4):268–274.
[40] Mathur MB, Ding P, Riddell CA, et al. Website and r package for computing e-values. Epidemiology. 2018;29(5):e45. (Cambridge, Mass).
[41] Dokubo EK, Shiraishi RW, Young PW, et al. Awareness of hiv status, prevention knowledge and condom use among people living with hiv in mozambique. PLoS ONE. 2014;9(9):Article ID e106760.
[42] Fedor TM, Kohler HP, Behrman JR. The impact of married individuals learning hiv status in malawi: divorce, number of sexual partners, and condom use with spouses. Demography. 2015;52(1):259–280.
[43] Kalichman SC, Ntseane D, Nthomang K, et al. Recent multiple sexual partners and hiv transmission risks among people living with hiv/aids in botswana. Sex Transm Infect. 2007;83(5):371–375.
[44] Commission MNA, et al. National hiv prevention strategy, 2015–2020. In: Lilongwe, editor. Malawi: Commission NA; 2014.
[45] Exavery A, Kanté AM, Tani K, et al. Sociodemographic drivers of multiple sexual partnerships among women in three rural districts of tanzania. HIV/AIDS. 2015;7:105. (Auckland, NZ).
[46] Onoya D, Zuma K, Zungu N, et al. Determinants of multiple sexual partnerships in south africa. J Public Health (Bangkok). 2015;37(1):97–106.
[47] Rakgoasi S. Association between awareness of hiv status and hiv risk behaviour among young people in botswana. South Afr J Demogr. 2018;18(1):171–219.
[48] Mishra VK, Bignami-Van Assche S. Concurrent sexual partnerships and hiv infection: evidence from national population-based surveys. Macro International, Incorporated; 2009.
[49] Mtenga SM, Pfeiffer C, Tanner M, et al. Linking gender, extramarital affairs, and hiv: a mixed methods study on contextual determinants of extramarital affairs in rural tanzania. AIDS Res Ther. 2018;15(1):1–17.
[50] Brown NR, Schweickart O, Sinclair RC, et al. Multiple partners, multiple strategies, multiple causes: Understanding why men and women provide discrepant sex partner reports. In: Cognitive development from a strategy perspective. Routledge; 2017. p. 43–60.

[51] King BM. The influence of social desirability on sexual behavior surveys: a review. Arch Sex Behav. 2022;51:1495–1501.
[52] Suarez D, Borràs R, Basagaña X. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. Epidemiology. 2011;22(4):586–588.
[53] Lusivika-Nzinga C, Selinger-Leneman H, Grabar S, et al. Performance of the marginal structural cox model for estimating individual and joined effects of treatments given in combination. BMC Med Res Methodol. 2017;17(1):1–11.
[54] Yang S, Eaton CB, McAlindon TE, et al. Effects of glucosamine and chondroitin supplementation on knee osteoarthritis: an analysis with marginal structural models. Arthritis Rheumatol. 2015;67(3):714–723.
[55] Pega F, Blakely T, Glymour MM, et al. Using marginal structural modeling to estimate the cumulative impact of an unconditional tax credit on self-rated health. Am J Epidemiol. 2016;183(4):315–324.
[56] Tsai AC, Weiser SD, Petersen ML, et al. A marginal structural model to estimate the causal effect of antidepressant medication treatment on viral suppression among homeless and marginally housed persons with hiv. Arch Gen Psychiatry. 2010;67(12):1282–1290.
[57] Romero-Estudillo E, González-Jiménez E, Mesa-Franco MC, et al. Gender-based differences in the high-risk sexual behaviours of young people aged 15-29 in melilla (spain): a cross-sectional study. BMC Public Health. 2014;14(1):1–9.
[58] Wirth K, Bennett K, Chakalisa U, et al. Self-reported risky sexual practices among adolescents and young adults in botswana. South Afr J HIV Med. 2019;20(1):1–8.
[59] Fenton KA, Johnson AM, McManus S, et al. Measuring sexual behaviour: methodological challenges in survey research. Sex Transm Infect. 2001;77(2):84–92.
[60] Durant LE, Carey MP. Reliability of retrospective self-reports of sexual and nonsexual health behaviors among women. J Sex Marital Ther. 2002;28(4):331–338.
[61] Michael H, Cui Y, Lorch S, et al. Instrumental variable estimation of marginal structural mean models for time-varying treatment; 2020. Available from: https://arxiv.org/abs/2004.11769.
[62] Cui Y, Michael H, Tanser F, et al. Instrumental variable estimation of the marginal structural Cox model for time-varying treatments. Biometrika. 2021;103(1): 1–18.

Appendices

Appendix 1. Distribution of loss-to-follow-up across the confounders

Table A1. Distribution of characteristics by loss-to-follow-up in the MLSFH in 2006.

	Loss-to-follow-up (LTFU)		
	Yes	**No**	**p-value**
	HIV awareness		
HIV positivity	72 (90.0)	1306 (94.8%)	
HIV-negativity	8 (10.0%)	72 (5.2%)	0.068
	Gender		
Female	54 (67.5)	905 (65.7%)	
Male	26 (32.5%)	474 (34.3%)	0.738
	Age category		
14-44 years	61 (76.3%)	964 (69.9%)	
45 years above	19 (23.8%)	414 (30.0%)	0.231
	Wealth		
Poor	24 (30%)	503 (36.5%)	
Medium/rich	56 (70.0%)	875 (63.5%)	0.239
	Education		
None	15 (18.8%)	321 (23.3%)	
Primary	57 (71.3%)	912 (66.2%)	
Secondary/Tertiary	8 (10.0%)	145 (10.5%)	0.610
	Religion		
Christian	47 (58.8%)	789 (57.3%)	
Muslim or other	33 (41.3%)	589 (42.7%)	0.097
	Region		
North	44 (55.0%)	281 (20.4%)	
Central	17 (21.3%)	531 (38.5%)	
South	19 (23.8%)	566 (41.1%)	< 0.001

Appendix 2. Odds ratio results for the MSM with unstabilized weights

Table A2. Estimates of the effect of awareness of HIV positivity on risky sexual behavior using unstabilized weights.

Awareness of HIV status	**Odds ratio**	**95% CI**	**p-value**
	(a) Effect on condom use		
HIV negativity	1.00		
HIV positivity	2.87	(1.67, 4.95)	< 0.001
	(b) Effect on multiple sexual partners		
HIV negativity	1.00		
HIV positivity	0.94	(0.48, 1.84)	0.85

95% Confidence Interval

Appendix 3. Results on the sensitivity analysis

Table A3. E-value to assess robustness to unmeasured confounding for the causal effect of HIV positivity awareness on condom use and multiple partners.

Outcome	E-value for point estimate	E-value for LL of CI
	Condom use	
HIV positivity	3.87	2.98
	Two or more partners	
HIV positivity	1.31	1.00

LL – Lower limit
CI – Confidence Interval
E-value parameters calculated as a function of the adjusted odds ratio (OR) for the effect of HIV positivity awareness on the outcomes