Causal Effect of Maternal HIV on Child Health Outcomes in the Context of Prevention of Mother to Child HIV Transmission (PMTCT) in Malawi

Halima S Twabi¹*, Samuel O Manda¹,²,⁴ and Dylan S Small³
Abstract

Background: The World Health Organisation recommends HIV-infected pregnant mothers to receive anti-retroviral drugs and to exclusively breastfeeding for the first six months within the Prevention of Mother to Child HIV Transmission (PMTCT) programme. Evidence has shown that maternal HIV infection has adverse effects on child birth weight. However, the effect of anti-retroviral therapy (ART) on pregnancy outcome is not conclusive. Ascertainment causality of these associations in high HIV epidemic and ART up-to-date populations remain largely unexamined and untested, and if confirmed would help policy makers to improve implementation and sensitization of PMTCT program. However, assessing causal effect has been limited due to ethical concerns if randomized controlled trials are appropriate. We aim to estimate the causal effect of maternal HIV on birth weight and the causal effect of knowledge of HIV status on exclusive breastfeeding while ascertaining the mediating effect of ART using observational data.

Methods: Data on over 18,000 and 16,000 children still alive and born within five years of the 2010 and 2015-16 Malawi Demographic and Health Surveys (MDHS) were analyzed. A set of methods for confounder balance namely, the 1:1 nearest neighbour (NN) matching, matching on the propensity score (PS) and inverse weighting the propensity score (PS) were used.

Results: Before matching, place of residence, region, anaemia level of a mother and age were statistically different between HIV-infected and HIV-uninfected mothers for both the 2010 and 2015-16 MDHS data while education and wealth were significantly different for the 2010 and 2015-16 data, respectively. After matching, the selected confounding variables were distributed similarly between HIV-infection status. Using the 2010 data, maternal HIV infection had a negative effect on infant birth weight (-142.2 (-281.4, -26.4); -149.0 (-324.2, -26.2) and -154.3 (-295.4, -13.3), under NN, PS matching and PS inverse weighting respectively. However, for the 2015-16 data, mothers who were infected with HIV gave birth to infants who significantly weighted more than infants more from HIV-uninfected mothers 125.5 (29.0, 219.9); 144.7 (40.4, 249.1) and 115.4 (11.6, 219.2), under NN, PS matching and PS inverse weighting respectively. Uptake of ART had no bearing on the adverse effect of HIV infection on birth weight. Maternal knowledge of HIV status was found to be associated with their behaviour concerning exclusive breastfeeding for the 2015-16 data. A sensitivity analysis suggested that causal effect estimates and p-values would change with a small bias due to unmeasured confounders.

Conclusion: We have found conflicting evidence on the association between maternal HIV infection and birth weight. The adverse association between maternal HIV infection and child birth weight found for the 2010 data could more likely be causal. However, the increased birth weight among infants born to HIV infected mothers in 2015-16 may show the recent successes of policies and interventions within the PMTCT program in Malawi. However, the purported mediating effect of anti-retroviral medication on the association between HIV infection and birth weight was not found. Prenatal care interventions should continue to be supported among ANC clinics in Malawi.

Keywords: Propensity score; Confounders; Maternal HIV; Inverse weighting; Causal effects
Background

The Human Immunodeficiency Virus (HIV) has been one of the largest public health challenges especially in low and middle income countries (LMICs). The World Health Organisation (WHO) estimates that more than 90% of paediatric HIV/AIDS infection is accounted for by Mother-to-child transmission (MTCT). The main consequences of maternal HIV infection to fetal health are mother-to-child transmission (MTCT) and the potentially increased risk of adverse pregnancy outcomes such as low birth weight. It is suggested that low birth weight is associated with perinatal morbidity and mortality but also with long-term adverse health outcomes. Identifying important and modifiable factors that causally influence these child birth outcomes will improve understanding of the pathways underlying the associations, thereby provide clues for possible prevention and intervention with the PMTCT programme.

Malawi is one of the countries that has made efforts to implement the Prevention of Mother to Child HIV Transmission (PMTCT) program. An effective PMTCT programme requires HIV-infected mothers and their infants to have access to and take up a cascade of interventions, of which some of them include enrolment of HIV-infected pregnant women and their families into anti-retroviral treatment (ART), safe childbirth practices and appropriate infant feeding.

The aim of PMTCT programmes is to reduce vertical transmission from mother to child and the program has been successful in reducing vertical HIV transmission, but implications of uptake of PMTCT interventions on adverse pregnancy outcomes and appropriate infant feeding have been reported. There has been conflicting evidence in literature on the association of maternal HIV on adverse pregnancy outcomes such as low birth weight. Some studies have claimed that ART among HIV-infected mothers is associated with adverse outcomes such as low birth weight (LBW), intrauterine growth restriction (IUGR), preterm delivery (PTD) and still births. Other studies have attributed these adverse outcomes to maternal HIV. Even though the mechanism on how HIV infection causes LBW remains unknown, it is believed that complications related to HIV infection impairs placental functions resulting in LBW. Lambart et al. suggests that prenatal care and antiretroviral therapy may reduce adverse pregnancy outcomes such as LBW.

Malawi is one of the highest HIV-affected countries in the world with 8.8% of the adult population (aged 15-49) living with HIV with a high prevalence rate among women as compared to men (10.8% to 6.4%). Breastfeeding being an important feeding practice, determines a child’s nutritional status, hence influencing the child’s growth and development.
In preventing mother to child transmission (MTCT), WHO recommends HIV-infected mothers from resource limited settings to receive ART and practice exclusive breastfeeding for the first 6 months postpartum. HIV transmission through breastfeeding can be reduced if HIV-positive women breastfeed exclusively for six months rather than practising mixed feeding. Evidence suggests that HIV-infected mothers who are on ART and mix-feed may have a higher rate of transmission than mothers who exclusively breastfeed and are on ART. Exclusive breastfeeding (EBF) is one of the most effective global public health intervention for child survival \[16, 17\]. Breast milk has nutrients and antibodies that are essential for child growth and a defensive mechanisms against antigens, respectively. Therefore there is a need to emphasise the importance of exclusive breastfeeding.

In an effort to help policy makers ensure that there is effective uptake of PMTCT guidelines such as accessing and utilising ANC services among HIV-infected pregnant mothers, and promoting and supporting appropriate infant feeding among HIV-infected mothers for child survival and prevention of infectious diseases, we aim to estimate the causal effect of maternal HIV on child birth weight while ascertaining the mediating effect of ART status on this effect and the causal effect of knowledge of a mothers HIV status on the practice of exclusive breastfeeding.

In order to determine that an observed association between exposure and outcome is more likely to be causal, the ideal methodology would be to conduct a randomised control trial (RCT) but it would be unethical to randomise mothers into HIV-infected and HIV-uninfected groups. There is readily available observational data, such as population and household-based surveys that provide most of the maternal and child health indicators, but confounding variables that are associated

The overall context of the PMTCT program is to mitigate the adverse effect of HIV among mothers on maternal and child health outcomes.

A major recommendation for PMTCT guidelines for mothers living with HIV, is exclusive breastfeeding, where infants receive breast milk only with an exception of appropriate medication.

Several studies have demonstrated that post-natal mother to child transmission through breastfeeding is less likely in infants who are exclusively breastfed than those given mixed feeds in the first six months of life.

Mothers who are HIV infected are exposed to constant counselling on good nutrition and infant feeding through antenatal care visits or mother support groups. Previous studies have attributed these counselling programs to HIV infected mothers being more likely to exclusively breastfeed their infants compared to mothers who were HIV uninfected.

Through studies have shown a significant association of maternal HIV on birth weight and exclusive breastfeeding, however, there is a need to establish a causal relationship and to measure the uptake of PMTCT guidelines by HIV-infected mothers in Malawi, especially on safe childbirth practices and appropriate infant feeding.
with both HIV-infection and the outcome is of concern. While we cannot control for unmeasured confounding, we use state of the art methods to control for measured confounding. [4] Rosenbaum and Rubin [18], [5] developed Propensity Scores (PS) to control for confounding bias. Propensity scores aim to emulate the randomization procedure of RCTs, as the distribution of measured confounders between exposure groups is made similar by using the PS. Some studies [19, 20, 21, 18, 22, 23, 24] have looked at basic properties of the propensity score in balancing difference between exposure groups based on subjects covariates, and others [25, 26, 24] have looked at a comparison of methods of balancing based on the propensity score methods with most studies concluding that inverse probability weighting is the best way to balance.

Therefore using survey data that is readily available and nationally representative of Malawi, we aim to establish the causal effect of maternal HIV on infant birth weight and causal effect of knowledge of HIV status on exclusive breastfeeding. We use three methods; Nearest neighbour matching, Propensity score (PS) matching and Inverse weighting on the PS that are commonly used to ascertain causal effect from observational data and we define assumptions needed to validate causal inference from cross-sectional studies, that has rarely been used.

Methods

Statistical Analysis

In order to assess causality of maternal HIV (exposure) on infant birth weight and causal effect of knowledge of mothers HIV status on exclusive breastfeeding from observational data such as MDHS, the distribution of covariates that act as confounders must be the same between HIV-infected and HIV-uninfected mothers so that the effect on birth weight should be attributed to HIV status only. Methods used to control for confounding aim to balance the differences of observed covariates between HIV-infected and HIV-uninfected mothers hence estimating an unbiased causal effect. One method, apart from the well known covariate adjustment, is matching [27]. We estimated the average treatment effect, \( ATE = E(Y_i(1) - Y_i(0)) \) on birth weight and exclusive breastfeeding by comparing the outcomes between HIV-infected and HIV-uninfected mothers across variables that were considered as possible confounders using a combination of exact and NN
matching, propensity score matching and Inverse probability weighting based on the propensity score. We present a brief description of three matching methods, a combination of exact and nearest neighbour matching and the propensity score matching. We also present the method of inverse probability weighting on the PS and assumptions that validate using it to estimate the causal effect from cross-sectional studies.

**Exact and Nearest Neighbour (NN) matching**

Exact matching involves finding a match on observed characteristics for each HIV-infected (exposed) subject \( i \) to a set of unmatched HIV uninfected (unexposed) subjects \( j \) such that \( X_i = X_j \) [22]. Once the matching procedure is satisfied, it would demonstrate that the distribution of covariate vector for \( X_i \), \( f(X_i) \) is similar to the distribution of the covariate vector \( X_j \), \( f(X_i) \). Austin, [28] in his paper states that exact matching on important covariates is ideal whenever possible to ensure a high quality comparison group. The main purpose of exact matching was to obtain HIV uninfected mothers \( j \) who were similar to HIV infected mothers \( i \) based on their measured covariates such that \( X_i = X_j = x \), where \( x \) is the matched vector. The matches were obtained when the distance denoted as \( D_{ij} \) was as follows:

\[
D_{ij} = \begin{cases} 
0 & \text{if } X_i = X_j \\
\infty & \text{if } X_i \neq X_j 
\end{cases}
\]

The distance would be equivalent to zero when the covariates between HIV infected and HIV uninfected are similar. As \( X \) increases dimensionality, matching exactly on the covariates tends to be difficult [29]. In this case, methods that involve matching on distance metrics are appropriate.

The ATE was estimated using Nearest neighbour matching [27], a common distance metric, by comparing birth weight between HIV-infected and HIV-uninfected mothers using nearest neighbour matching across a set of confounding variables. [27] The simplest nearest neighbour matching method uses a “greedy” algorithm, which cycles through each exposed subject one at a time, selecting the available unexposed subject with the smallest distance to the exposed subject [29]. The 1 : 3 nearest neighbour matching selected for each HIV-infected mother \( i \), three corresponding HIV-uninfected mothers \( j \) with the smallest distance between them. The NN matching was done using the Mahalanobis distance metric, mathematically defined as:

\[
D_{ij}(X_i, X_j) = (X_i - X_j)^T \sum^{-1}(X_i - X_j)
\]

where \( X_i \) is the vector of observed covariates for HIV infected mothers and \( X_j \) is the vector of observed covariates for HIV uninfected mothers and \( \sum^{-1} \) is the

---

[c1] HT: or exclusive breastfed subject
[c2] HT: or non-exclusive breastfed
[c3] HT: or non-exclusive breastfed infants
[c4] HT: or exclusive breastfed infants
[c5] HT: or exclusive breastfed
[c6] HT: or non-exclusive breastfed
[c7] HT: The methods selects \( k \) matched HIV uninfected (unexposed) subjects for each HIV-infected (exposed) subject, most often for \( k = 1 \).
[c8] HT: non-HIV
covariance matrix between the variables. Gu and Rosenbaum [30] did stated that use of Mahalanobis distance (with or without calipers) is best when there are few covariates being adjusted as we have in this study. In an effort to obtain perfect matches we performed a combination of 1 : 3 nearest neighbour matching and exact matching. The NN matching without replacement using Mahalanobis distance score was done for variables region, residence, anaemia and wealth. Exact match was done on variables region and age for exposure HIV [c9].

Matching based on the Propensity Score (PS)

Rosenbaum et al. [18, 31] introduced an alternative to the traditional matching between exposed and unexposed on covariates called a propensity score. The mathematical expression for the propensity score [18] is as follows:

\[ e(x_i) = P(HIV = 1 | x_i) \]  \hspace{1cm} (1)

A score for the probability that a subject is HIV-infected conditional on observed covariates is obtained. The propensity score provides a single measure of influence of confounders on exposure assignment. The propensity score is a balancing score and Rosenbaum [18] showed that this balancing score, under the strong ignorability assumption can produce unbiased average treatment effects. This propensity score is then used to balance between HIV-infected or HIV-uninfected groups either through matching, stratification [19], covariate adjustment or inverse weighting [25, 21, 20, 32]. The advantage of the propensity score is that it can consider more matching factors and improve efficiency of a study [33].

The propensity score was estimated using a binary logistic regression with exposure (HIV status= 1 or 0) as an outcome variable against potential confounders. Other studies have estimated the p-scores using boosted regression methods [34]. Variables which were considered as important for matching were included as covariates. [c3] We considered these covariates for effect of maternal HIV on birth weight; region, place of residence, wealth, age, maternal anaemia, smoking status of a mother and education level. For the effect of maternal HIV on exclusive breastfeeding we consider covariates; region, place of residence, wealth, age and wealth, whether a child had diarrhoea, fever. We further looked at the mediating effect of ART status on effect of maternal HIV on birth weight. In this study, the propensity score was estimated for outcome HIV status, using a multiple logistic regression model. The model was as follows

\[
P(HIV \text{ Status } = 1 \mid X_1) = \log \left( \frac{\pi(x)}{1-\pi(x)} \right) = \beta_0 + \beta_1 \text{ Region} + \beta_2 \text{ age} + \beta_3 \text{ residence} + \beta_4 \text{ Education} + \beta_5 \text{ Wealth} + \beta_6 \text{ fever} + \beta_7 \text{ anaemia} + \beta_8 \text{ fever} + \beta_9 \text{ diarrhoea} + \beta_{10} \text{ smoking}
\]
where the vector of parameters $\beta$ are estimated based on the whole sample. $\pi(x)$ is the probability of a mother being HIV-infected ($Y = 1$)

Matching on the propensity score compared the distance or difference between a score for an HIV-infected and a HIV-uninfected mother. Mathematically we can present it as $D_{ij} = |e_i - e_j|$, $i \neq j$. Where $e_i$ is the propensity score for HIV-infected mothers and $e_j$ is the PS for HIV-uninfected mothers. The smaller the difference between the scores, the more similar they are, so there is a pair matching [25, 21]. Therefore from the original sample, just a part of it would be kept for the final analysis. We performed a 1:1 matching by using the calipers matching, which matches each HIV-infected mother and a HIV-uninfected mother who has similar propensity scores, within a caliper of 0.01 for exposure HIV status. The matching is done within a specific region were the exposed and unexposed subjects would be matched using the propensity scores. [\textsuperscript{c3}]HIV-uninfected mothers with p-scores outside that region were pruned. Although matching with replacement is more advantageous as it would allow HIV-uninfected (unexposed) mothers that look similar to many of the HIV-infected (exposed) individuals to be used multiple times [29], we used matching without replacement.

Causal Effect Estimation and Post-Matching analysis

Statistical comparison among the matched samples

Matching by the PS creates a balanced dataset allowing a direct comparison of baseline covariates between HIV-infected and HIV-uninfected mothers [18]. To estimate the ATE from the matched sample using the propensity score, a comparison we compare birth weight outcome between HIV-infected and HIV-uninfected mothers on the matched sample. Austin [35] recommended the use of statistical methods that take into account the matched sample. The estimates obtained after matching are valid under the assumption of strong ignorability (no unmeasured confounders). Testing for balance among the categorical covariates between the exposure groups were done using a standardized bias. The standardized bias compared the prevalence of the covariate in the exposed and non-exposed subjects [35]. Covariates were balanced if the |bias| < 5%. We estimated the difference in means and proportions between HIV infected and HIV-uninfected mothers, in the propensity score matched sample across covariates. We separated the full sample (HIV-infected mothers) into five quintiles defined by their propensity scores. For the covariates, We compared frequencies before and after adjusting for propensity score quintile among HIV-infected and HIV-uninfected mothers. Specifically, we compared the p-values for HIV status and after adjustment for propensity score quintile with the p-values for HIV status before adjustment to determine whether balance on the covariates was achieved.

[c1] non-HIV
[c2] non-HIV
[c3] non-HIV
[c4] non-HIV
[c5] HT: Text added.
[c6] HT: HIV mothers
Inverse Probability Weighting (IPW) based on the PS

One of the challenges with matching on the propensity scores is that the number of matched pair sample is small as compared to the original sample. This means a large part of the sample would not be used to estimate the causal effect. An alternative method to matching is the Inverse probability weighting (IPW) based on the PS. The PS weighting aims to re-weight HIV-infected mothers and HIV-uninfected mothers to make them representative of the population of interest [19]. It is used to balance confounding between exposures by creating a pseudo-population that is free from confounders by assigning weights that have been obtained by adjusting for the measured confounders. Studies [25, 28] have shown that inverse weighting on the PS performed better than matching on the PS. 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

$$w_i = \frac{1}{1-e(x_i)}.$$  

To estimate an average treatment effect based on the weights, the inverse probability weights can be defined with respect to the exposure variable as:

$$w_i = \frac{A_i}{e(x_i)} + \frac{(1 - A_i)}{1 - e(x_i)}, A_i = 0, 1$$  \hfill (2)

The IPW controls the influence of subjects by weighting their responses based on their propensity score [34]. We estimated the IPW as a weight on the PS using the p-scores obtained for mothers HIV status on confounders, place of residence, education and wealth quintile, age and region, smoking status of a mother, malaria during pregnancy and maternal anaemia, whether a child had a fever, malaria. To estimate average treatment effect using IPW, a weighted logistic regression model was fit using the inverse probability weights for the outcome variables birth weight and exclusive breastfeeding on the exposure, mothers HIV status.

Assumptions to Validate Causal Effect

Proper defining of conditions needed for valid estimation of causal effects in cross-sectional studies, can enable one to estimate prevalence causal estimates in cross-sectional studies as explained by Flanders et al [36]. He states that if certain conditions (assumptions) are met, then a prevalence odds ratio can provide a valid causal estimate in cross-sectional studies. Flanders et al, [36] states that in order to estimate causal effect using the prevalence odds ratio or prevalence difference from cross-sectional studies, survey participants must be randomly selected from the survey population $P_1$, the target population $P_0$ and age at exposure must also be defined.

A natural estimator of the population average causal effect of exposure (E) on presence of an outcome (Y) in population $P_0$ is the observed difference:

$$\hat{cPD} = \sum_{(1,a_1)} - \sum_{(0,a_1)}$$  \hfill (3)

[c1] We estimated the IPW as a weight on the PS using the p-scores obtained for mothers HIV status and exclusive breastfeeding on confounders, place of residence, education and wealth quintile, age and region.
Where $cPD$ refers to the causal prevalence difference and $\bar{Y}$ represents the average of the observed prevalence of outcome $Y$ at age $a = a_i$ in HIV status group $A_i$ ($A = 1$ if HIV-infected, 0 otherwise). We defined the following assumptions and validated as illustrated by Flanders et al. [36] which allowed us to estimate the causal prevalence odds ratio of maternal HIV on exclusive breastfeeding.

i. The baseline target population must consist of subjects who are alive and emanate from a larger population that was subjected to exposure at a particular time. For the exposure-risk factor association conditions, the larger population for this study was defined as HIV-infected mother or who exclusively breastfed their infant within 0-6 months of age (exposure) five years before the survey. The baseline target population was defined as all women who were alive and were either HIV-infected or HIV-uninfected. It also consisted of babies who had been exclusively breastfed and were alive at the time of the survey.

ii. Exchangeability must be met. This is where the larger target population $P_0$ should be defined. This is done by identifying the causal relationships between exposure and other causes of the outcome variable which are known as confounders. The study achieved this assumption by balancing the differences between HIV-infected and HIV-uninfected mothers among confounders that we felt affected both the HIV status and birth weight. The balancing was done by using the propensity score matching and inverse probability weighting.

iii. Independence of subjects was assumed (no interdependence between subjects in a cluster) The MDHS is collected using cluster sampling, since individuals are randomly selected within enumeration areas which are clusters, the Stable Unit Treatment Value Assumption (SUTVA) was defined under condition that there was no interdependence between clusters and because subjects/units were randomly selected, selection bias was reduced, as such no interdependence between subjects and within a cluster.

Sensitivity Analysis

Matching on the propensity score with the assumption of strong ignorability implies that we assume there are no unobserved confounders that would affect our causal effect estimate. It is important to measure to what level would unobserved confounders affect our causal estimates, since the strong ignorability assumption cannot be assessed. The [c1] Malawi Demographic and Health Survey (MDHS) data has numerous variables that are likely to influence birth weight and exclusive breastfeeding. We controlled for variables that showed significant difference in distribution between HIV-infected and HIV-uninfected groups. Other variables such as, clinical factors, mothers diet, place of delivery, age of infant at birth, fetal height were not observed in DHS and would also affect our response variable. In order to assess if our causal estimates obtained after matching were sensitive to hidden bias, we used a sensitivity analysis developed by Rosenbaum et al. [37].

The sensitivity analysis method for matched data provides a parameter that deals with magnitude of hidden bias that would need to be present to explain the association observed [38]. Rosenbaum’s [38] method of sensitivity analysis relies on the
sensitivity of the parameter \( \Gamma \) which measures the degree of departure from a study that is free of bias. Two mothers with the same measured characteristic may differ in the odds of being HIV-infected or HIV-uninfected by at most a factor of \( \Gamma \). If there are no unmeasured confounders then \( \Gamma = 1 \), but if there are unmeasured confounders \( \Gamma = 2 \), then these two mothers (HIV-infected and with the same observed characteristics (place of residence, education status, region, age [c2] anaemia level of a mother, whether a child had a fever, whether the child had diarrhoea) may differ in the odds of being HIV-infected by as much as a factor of \( 2 \) [39].

We obtained upper and lower bounds on the p-value for values of \( \Gamma \) using the following formula;

\[
\text{upper bound } p - \text{value} = \sum_{a} \binom{T}{a} (p^+)^a (1 - p^+)^{T-a}
\]

where \( p^+ = \frac{\Gamma}{1+\Gamma} \). Upper and lower bounds on the Hodges-Lehmann test for outcome birth weight, setting the maximum value of \( \Gamma \) to 1.3 with 0.05 increments was obtained. The rbounds package in R was used to obtain the upper and lower bounds on the p-value and point estimate.

**Study Data**

The data from this study was obtained from the 2010 and 2015-16 Malawi Demographic and Health Survey. These are nationally representative household survey that provide data for a wide range of population, health, and nutrition indicators focused on under five year old children in Malawi. [c1] The 2015-16 data provided up-to-date information on current HIV trends, child health and nutrition. The 2010 data provided us information of ART status confounding on the effect of maternal HIV on birth weight and exclusive breastfeeding from a PMTCT context. Parents of infants signed informed consent. The sampling design according to [c2] 2015/16 MDHS report [40] was a two stage cluster design with stratification. The primary sampling units were the enumeration areas (EAs) and the secondary sampling units were the households. EAs were stratified in terms of rural and urban. A total of 849 EAs were sampled with 158 in urban areas and 691 in rural areas. A total representative sample of [c3] 26564 households was selected and [c4] 26361 households were considered to be occupied in the [c5] 2015/16 MDHS. Data collection was by questionnaires. There were three types of questionnaires, woman, man and household questionnaire through face-to-face interviews. [c6] 26361 households were successfully interviewed, yielding a response rate of [c7] 99.2%. Eligible women were [c8] 25146 and [c9] 24562 were successfully interviewed, yielding a response rate of 97%. [c10] A sub-

---

[c2] HT: Text added.
[c4] HT: Text added.
[c6] HT: Text added.
sample of 7924 women were tested for HIV using an ELISA test. The data set used in the analysis was child record data set which was based on woman and household questionnaires.

The study analyzed data on over 18,000 and 16,000 children still alive and born within five years of the 2010 and 2015-16 Malawi Demographic and Health Surveys. The outcome variables of interest generated were: Birth weight (grams) and exclusive breastfeeding (0 = Non Exclusive breastfed 1 = Exclusive breastfed). The exposure variable of interest was Mothers HIV status (0 = HIV-uninfected, 1 = HIV-infected) and ART status among HIV-infected mothers (0 = Not on ARVs, 1 = On ARVs) using the 2010 MDHS. The confounders in the study were identified as region, place of residence, wealth quintile, Mothers level of education, age of mother, anemia level of a mother, mothers smoking status whether a child had a fever, whether the child had diarrhoea. Causal effect estimates were obtained for the effect of maternal HIV on birth weight and on exclusive breastfeeding.

Results
A summary of distribution of mothers’ baseline characteristics by mothers HIV status are shown in Table 1. Valid data was only available for 13987 and 11782 alive under-five children for the causal effect of HIV on birth weight and 1599 and 1465 still alive children aged 0-6 months to study the causal effect of HIV status knowledge on exclusive breastfeeding for the 2015-16 and 2010 MDHS data, respectively. Out of the 13987 children, 7912 children had mothers who knew their HIV status. 7033 (88.9%) of the mothers were HIV-uninfected and 879 (11.1%) were HIV-infected. Out of the 7912 mothers with known HIV status, 3694 were from the rural area with 291 (7.8%) HIV-infected. Out of 2890 of the mothers who had primary education 253 (9%) were HIV-infected. 1888 of the mothers were poor with 156 (8.3%) being HIV-infected.

Of the 1599 infants, 877 (55%) were exclusively breastfed and 722 (45%) were not exclusively breastfed. Only 497 mothers of the infants knew their HIV status. 449 (90.3%) were HIV-uninfected and 48 (9.7%) were HIV-infected. Out of the 48 of the mothers who were HIV-infected, 25 (52.1%) exclusively breastfed their infants.

Text added.

Text added.

Analysis was done on 2756 child records.

demographic and exclusive breastfeeding practices

Text added.
Using the 2010 DHS, the analysis was done on 11782 alive children. A total of 10151 infants of whom 96% (10151) of their mothers were HIV-uninfected and 427 mothers were HIV-uninfected. Of the 427 HIV-infected mothers, 216 (50.6%) were on ART. A majority of the children were from the rural area (9269) and 7270 of the mothers had a primary education. 5111 of the mothers were aged 25-34 years old and of this sample 250(4.9%) of the mothers were HIV-infected.

There were a total 1465 infants who were aged 0-6 months in the 2010 DHS. 922 (62.9%) were exclusively breastfed and 543(37.1%) were not exclusively breastfed. Only 1369 mothers of the infants knew their HIV status. 63 of the 1369 were HIV-infected and only 37(58.7%) of the 63 exclusively breastfed their infants.

A total of 2708 infants were breastfed with 917(33.9%) being exclusively breastfed. A majority of the mothers were from the rural area representing 87.9% of the sample. The majority were poor and 69.4% of the mothers had a primary education. Among the 87.9% of women living in the rural area, 12.4% exclusively breastfed their infants. Of the 69.4% of the women who had primary education, 634 (69.1%) exclusively breastfed their infants. Before matching, participants demographic characteristic age varied among exclusively breastfed and non-exclusively breastfed infants as presented in Table 1b. Exclusive breastfeeding was higher among infants whose mother were aged 20-24 years old.
Table 1: Distribution of Confounders between HIV-infected and HIV-uninfected before matching for MDHS 2010 and 2015-16

| Characteristic | Mothers HIV status (2010) | Mothers HIV Status (2015) |
|---------------|---------------------------|---------------------------|
|               | HIV-infected(%) | HIV-uninfected(%) | p-value | HIV-infected(%) | HIV-uninfected(%) | p-value |
| Overall       | 427(4)          | 10151(96)         |         | 903(8.7)       | 7173(91.3)       |         |
| Residence     |                |                   |         |                |                   |         |
| Urban         | 89(6.9)         | 1222(93.2)        | <0.001  | 124(15.4)      | 676(86)          | <0.001  |
| Rural         | 338(3.7)        | 8931(96.3)        |         | 291(7.8)       | 3401(92.1)       |         |
| Region        |                |                   |         |                |                   |         |
| Northern      | 38(1.7)         | 2267(98.4)        |         | 48(5.7)        | 794(94.3)        |         |
| Central       | 74(2.1)         | 3432(97.9)        |         | 94(6.1)        | 1459(93.9)       |         |
| Southern      | 315(6.6)        | 4452(93.4)        | <0.001  | 273(13.0)      | 1827(87.0)       | <0.001  |
| Age           |                |                   |         |                |                   |         |
| 15-24         | 58(1.6)         | 3562(98.4)        |         | 101(6.1)       | 1558(93.9)       |         |
| 25-34         | 250(4.9)        | 4861(95.1)        |         | 195(9.5)       | 1861(90.5)       |         |
| 35 above      | 119(6.4)        | 1728(93.6)        | <0.001  | 119(15.3)      | 661(84.7)        | <0.001  |
| Wealth        |                |                   |         |                |                   |         |
| Poor          | 161(3.9)        | 3950(96.1)        |         | 156(8.3)       | 1732(91.7)       |         |
| Middle        | 94(4.0)         | 2265(96)          |         | 69(8.1)        | 787(91.9)        |         |
| Rich          | 172(4.2)        | 3963(95.8)        | 0.815   | 190(10.9)      | 1561(89.1)       | 0.011   |
| Education     |                |                   |         |                |                   |         |
| None          | 68(5.3)         | 1205(94.7)        |         | 50(10.4)       | 432(89.6)        |         |
| Primary       | 292(4.0)        | 6978(96)          |         | 253(9.0)       | 2637(91)         |         |
| Secondary     | 65(3.3)         | 1881(96.7)        |         | 106(10.2)      | 937(89.8)        |         |
| Tertiary      | 67(97.8)        | 2(2.2)            | 0.032   | 6(7.5)         | 74(92.5)         | 0.407   |
| Anaemia       |                |                   |         |                |                   |         |
| No            | 96(3.8)         | 2414(96.2)        |         | 457(8.7)       | 4796(91.3)       |         |
| Yes           | 49(5.4)         | 851(94.6)         | 0.039   | 422(15.9)      | 2234(84.1)       | <0.001  |
| Smoking Status|                |                   |         |                |                   |         |
| No            |                |                   |         |                |                   |         |
| Yes           |                |                   |         |                |                   |         |
| Had fever last 2 weeks |               |                   |         |                |                   |         |
| No            | 277(4.0)        | 6593(96)          | 0.89    | 301(9.6)       | 2827(90.4)       | 0.31    |
| Yes           | 149(4.1)        | 3499(95.9)        |         | 111(8.5)       | 1203(91.6)       |         |
| Had Diarrhoea |                |                   |         |                |                   |         |
| No            | 247(4.0)        | 8276(96)          |         | 70(9.6)        | 659(90.4)        |         |
| Yes           | 79(4.2)         | 1809(95.8)        | 0.722   | 17(5)          | 219(94)          | 0.353   |

Comparison Before and After matching

Before matching, using the 2015-16 MDHS, participants demographic and health characteristics region, place of residence, age, anaemia level and wealth varied among HIV-infected and HIV-uninfected mothers as presented in table A1. For the 2010 MDHS, characteristics residence, region, age, education level and anaemia were different among HIV-infected and HIV-uninfected mothers as presented in table A1.

After PS matching on the covariates for exposure mothers HIV status, 116 matched pairs (HIV-infected and HIV-uninfected) for the 201 MDHS data and 467 matched pairs for the 2015-16 MDHS data were analysed for difference in baseline characteristics and no significant difference was observed between the HIV-infected and HIV-uninfected matched sample as presented in Table A1. Before matching the absolute value of standardised bias was greater than 10 across the covariates except for smoking status of a mother. After matching the standardised bias ranged from -0.4 to 0.5 with a high percentage reduction in bias as presented in [c1].

[c1] HT: Table 6 presents the distribution of covariates for the 917 matched pairs (exclusively breastfed and non-exclusively breastfed). There were no differences between the matched group for exposure exclusive breastfeeding.
Table 2. We stratified the matched observations into quintiles according to their propensity scores, we observe that within each quintile there is no significant difference between HIV infection status for the respective covariates as shown in tables A2 and A3. This demonstrates that balance has been achieved in terms of distribution of confounders between HIV-infected mothers and HIV-uninfected mothers in the matched sample.

|                      | 2010 DHS Unmatched Bias(µ) | 2010 DHS Matched Bias(µ) | 2015 MDHS Unmatched Bias(µ) | 2015 MDHS Matched Bias(µ) |
|----------------------|---------------------------|--------------------------|-----------------------------|--------------------------|
| Residence            | -30.8(<0.001)             | 5.4(0.677)               | -31.7(<0.001)               | -0.4(0.960)              |
| Wealth               | 14.8(0.076)               | 6.5(0.581)               | 13.7(0.008)                 | 0.1(0.99)                |
| Age                  | 39.2(<0.001)              | 0.2(0.988)               | 36.8(<0.001)                | -0.3(0.962)              |
| Anemia               | 16.9(0.039)               | 0.3(0.978)               | 39.3(<0.001)                | 0.5(0.945)               |
| Region               | 51.6(<0.001)              | 0.6(0.956)               | 39.7(<0.001)                | 0.4(0.947)               |
| Smoking              | 0.6(0.958)                | 0.4(0.963)               | 5.2(0.241)                  | 0.0(1.00)                |

Estimation of causal effects

We defined and reviewed conditions needed for valid estimation of causal effects in cross-sectional studies. The defined assumptions enables the estimation of the causal odds ratio and causal mean difference. The causal odds and difference were obtained based on the assumptions specified in the methods section. The average treatment effect was estimated as a difference in mean birth weight and as a comparison of odds of exclusive breastfeeding among HIV-infected and HIV-uninfected mothers.

Table 3 shows the mean difference in birth weight between infants born to HIV-infected or HIV-uninfected mothers and the causal odds of exclusive breastfeeding between HIV-infected and HIV-uninfected mothers for the 2015-16 MDHS data. The mean birth weight difference of children born between HIV-infected and HIV-uninfected mothers was 64.5g before matching. After nearest neighbour matching the mean birth weight difference between the two groups was 124.5g. For PS matching and Inverse probability weighting on the PS the mean difference of children’s birth weight was 144.7g and 115.4g respectively. The average maternal HIV effect on child birth weight was significant for the methods of NN matching, PS matching and PS weighting.

[c3]
The causal effect of maternal HIV on exclusive breastfeeding under no matching, nearest neighbour matching and PS weighting showed no significant difference in terms of exclusive breastfeeding practices between HIV-infected mothers and HIV-uninfected mothers. The ATE under PS matching provided a significant causal effect of knowledge of HIV status on exclusive breastfeeding. The results form the 2015-16 MDHS show that HIV-infected mothers have 9% and 18% lower odds to exclusively breastfeed than HIV-uninfected mothers using PS matching and PS weighting, respectively.\(^\text{[7]}\)\(^\text{[8]}\)

**Table 3**: Average Treatment Effect for Maternal HIV on Birth weight and Exclusive Breastfeeding

|                      | Child Health Outcomes 2015-16 | Child Health Outcomes 2010 |
|----------------------|------------------------------|-----------------------------|
|                      | Birth Weight | Exclusive breastfeeding | Birth Weight | Exclusive breastfeeding |
| Mean.Diff(95% CI)    | OR(95% CI)   | Mean.Diff(95% CI)        | OR(95% CI)   |
| Unmatched            | 64.5(-141.1, 12.1) | 1.03(0.49, 2.18) | -86.0(-161.2, -11.1) | 0.5(0.2, 1.3) |
| NN matching          | 124.5(29.0, 219.9) | 0.99(0.86, 1.14) | -142.2(-281.4, -26.4) | 0.9(0.7, 1.2) |
| PS matching          | 144.7(40.4, 249.1) | 0.91(0.84, 1.00) | -149(-324.2, -26.2) | 1.0(0.8, 1.4) |
| PS weighting         | 115.4(11.6, 219.2) | 0.82(0.68, 1.00) | -154.3(-295.4, -13.3) | 0.97(0.8, 1.2) |

**Table 3**\(^\text{[1]}\) also presents results for the estimates of the ATE of maternal HIV on birth weight and exclusive breastfeeding for the 2010 DHS data. The NN matching, PS matching and PS weighting showed a significant difference in mean birth weight between HIV-infected mothers and HIV-uninfected mothers. Children born to HIV-infected mothers had a 142.2g, 149g and 154.3g lower average birth weight as compared to infants born from HIV-uninfected mothers, respectively. There was no significant difference in terms of exclusive breastfeeding practices between HIV-infected and HIV-uninfected mothers for all methods.

**Table 4**\(^\text{[2]}\) presents results for the mediating effect of ART status for mothers who were HIV-infected on causal association of maternal HIV on birth weight and practice of exclusive breastfeeding. We observe that infants born from HIV-infected mothers who were on ART had a lower average birth weight as compared to infants born from HIV-infected mothers who were not on ART. The causal estimates for all the causal methods were not significant. Table 4 \(^\text{[3]}\) also presents mediating effect of ART status among HIV-infected mothers on exclusive breastfeeding. Despite the observed was not significant. This implies that assigning weights for each child created a pseudo population that had a wider difference in their birth weights between HIV-infected and HIV-uninfected mothers.

\(^{[6]}\)HT: Text added.

\(^{[7]}\)HT: The inverse weighting method produced an odds ratio that was slightly lower and had a narrower confidence interval (CI) unlike the CIs for after matching on the PS.

\(^{[8]}\)HT: After adjusting for inverse weights, HIV-infected mothers had 74% higher odds of exclusively breastfeeding their infants than mothers who were HIV-uninfected.

\(^{[1]}\)HT: Text added.

\(^{[2]}\)HT: Text added.

\(^{[3]}\)HT: Text added.
effect showing that HIV-infected mothers who are on ART are more likely to engage in exclusive breastfeeding, the causal effect for the three methods was not significant. ART status does not significantly impact exclusive breastfeeding practices among HIV-infected mothers.

Table 4: Mediating Effect of ART on Causal effect of maternal HIV on Birth weight and Exclusive Breastfeeding for 2010 DHS

| Child Health Outcomes | Birth Weight | Exclusive breastfeeding |
|-----------------------|--------------|-------------------------|
|                       | Mean.Diff(95% CI) | OR(95% CI)               |
| Unmatched             | [c2]-145.2(-302.1,11.6) | [c3]0.89(0.32, 2.44)     |
| NN matching           | [c4]-111.98(-407.183,1) | [c5]1.26(0.8, 2.0)       |
| PS Matching           | [c6]-162.1(-457.5,133.3) | [c7]1.24(0.92, 1.67)     |
| PS Weighting          | [c8]-129.7(-390.7,131.3) | [c9]1.40(0.9, 2.17)      |

Sensitivity Analysis

A sensitivity analysis was done after matching to observe if the causal estimates obtained were sensitive to bias from unmeasured confounders. Table 5a presents results on the upper and lower bounds on the Wilcoxon Signed Rank p-value and table 5b shows results of the upper and lower bounds on the Hodges-Lehmann point estimate for outcome birth weight. Based on the results, the p-value is slightly robust as it would require a small value of Γ for the upper and lower bounds to have a p-value of 0.05. The results inform us that with a small bias the causal effect of maternal on birth weight would be non-causal.

[c1] HT: Table 4 shows the average treatment effect, presented as a prevalence odds ratio of exclusive breastfeeding on stunting, wasting and underweight. Infants who were exclusively breastfed were 40%, 60% and 40% less likely to be wasted, stunted or underweight respectively using PS weighting. Using PS matching, exclusively breastfed infants were 60%, 50% and 30% less likely to be wasted, stunted and underweight, respectively. All the estimated causal effects were significant (p-values < 0.05).

[c10] HT: The nearest neighbour matching and PS matching produced a higher prevalence odds for wasting unlike for underweight.

[c11] HT: The estimate obtained after adjusting for the inverse weights was very different from the ones obtained under no match, NN matching and PS matching. The IPW had a wider confidence interval as compared to the other three methods. The estimates were slightly similar for the outcomes stunting and underweight but PS matching and PS weighting had narrower confidence intervals.
Table 5: Rosenbaum Sensitivity Analysis

(a) Rosenbaum Sensitivity Test for Wilcoxon Signed Rank P-Value and Hodges-Lehmann Point Estimate for 2010 MDHS

| Gamma | P-value Estimate |
|-------|------------------|
|       | P-value | Estimate |
|       | Gamma | Upper bound | Lower bound | Upper | Lower |
| 1     | 0.001 | 0.0007 | 400 | 400 |
| 1.1   | 0.002 | 0.0002 | 350 | 450 |
| 1.2   | 0.005 | 0.0001 | 300 | 500 |
| 1.3   | 0.01  | <0.0001 | 300 | 500 |
| 1.4   | 0.018 | <0.0001 | 250 | 550 |
| 1.5   | 0.029 | <0.0001 | 250 | 550 |

(b) Rosenbaum Sensitivity Test for Wilcoxon Signed Rank P-Value and Hodges-Lehmann Point Estimate for 2015-16 MDHS

| Gamma | P-value Estimate |
|-------|------------------|
|       | P-value | Estimate |
|       | Gamma | Lower bound | Upper bound | Lower | Upper |
| 1     | 0.0004 | 0.0004 | -83.3 | -83.3 |
| 1.1   | 0.0001 | 0.0016 | -133.3 | -133.3 |
| 1.2   | <0.0001 | 0.0052 | -150.1 | 150.1 |
| 1.3   | <0.0001 | 0.0135 | -166.7 | 166.7 |
| 1.4   | <0.0001 | 0.0294 | -183.3 | 183.3 |
| 1.5   | <0.0001 | 0.0556 | -200 | 200 |

The Hodges-Lehmann point estimate for the sign rank test presented in Table 5b provides additive effects due to HIV infection and is interpreted as the median differences between the HIV-infected and HIV-uninfected group. When \( \Gamma = 1 \), the upper bound estimate is -83.3, this suggests that the estimate is subject to change due to small bias caused by unobserved confounders as we note that at \( \Gamma = 1 \) the point estimate drastically changes.

Discussion

[c1] The study set out to assess the causal effect of maternal HIV and knowledge of a mother’s HIV status on infant birth weight and exclusive breastfeeding, respectively. The 2010 MDHS data provided a significant negative impact of maternal HIV on birth weight -142.2 (-281.4, -26.4), -149.0 (-324.2, -26.2), and -154.3 (-295.4, -13.3), under NN, PS matching and PS inverse weighting respectively. However, for the 2015-16 data, mothers who were infected with HIV gave birth to infants who significantly weighted more than infants born to HIV-uninfected mothers, 125.5 (29.0, 219.9), 144.7 (40.4, 249.1), and 115.4 (11.6, 219.2), under NN, PS matching and PS inverse weighting respectively. No significant effect of knowledge of HIV status on exclusive breastfeeding was shown for both the 2010 and 2015-16 MDHS.

[c2] Results from the 2010 MDHS have shown that children born from HIV-infected mothers had a lower birth weight as compared to children born from mothers who were HIV-uninfected. However, results from the 2015-16 MDHS data have shown that children born from HIV-infected mothers had a higher birth weight as compared to children born from HIV-uninfected mothers. Contrary to our findings for

[c1] HT: Text added.
[c2] HT: Text added.
the 2015-16 MDHS data and similar to the findings of the 2010 MDHS data, several studies found significant negative associations between maternal HIV status on child birth weight \cite{11, 13, 33}. Xiao et al., \cite{11} \footnote{HT: Text added.} through conducting a meta-analysis of cohort studies, found that HIV-infected women were at higher risk of having low birth weight babies compared with HIV-uninfected women. Other studies \cite{8, 12} \footnote{HT: Text added.} found that advanced-stage HIV disease among mothers was associated with higher risk of low birth weight among their infants. Findings from our study suggest an improvement in birth weight among children born to HIV-infected mothers in the 2015-16 MDHS. The results provide evidence that there have been major improvements in implementing PMTCT interventions by the ministry of Health and other key stakeholders. In July 2011, Malawi was one of the first country to implement Option B+ approach as an effort to deal with HIV/AIDS and its impact. This meant that all pregnant women living with HIV were offered ART for life irrespective of their CD4 count \cite{41}. \footnote{HT: Text added.} 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) known as "Chiponde" \cite{42}. \footnote{HT: Text added.} This has helped HIV-infected pregnant mothers to maintain a balanced nutritional status which has intern helped in child growth as compared to HIV-uninfected mothers who only receive iron supplements and folic acids without additional food supplements.

\footnote{HT: Text added.} ART status did not mediate the effect of maternal HIV infection on birth weight and the effect of knowledge of HIV status on exclusive breastfeeding practice using the 2010 MDHS. This implies that the observed causal association between maternal HIV infection and birth weight is direct and not influenced by ART status. ART status on it own should be considered as an exposure that would affect pregnancy outcomes rather than a mediating factor. Similar to our findings, Cotter et al., \cite{4} \footnote{HT: Text added.} using a prospective cohort study in the USA found that the risk of adverse pregnancy outcomes such as low birth weight attributable to ART was small. However they advised caution when deciding to administer a combination of ART drugs. Similarly, Marazzi et al., \cite{6} \footnote{HT: Text added.} in their retrospective cohort study done in Malawi and Mozambique, \footnote{HT: Text added.} found that low birth weight was not associated with ART duration and concluded that ART actually led to a reduction in adverse pregnancy outcomes such as low birth weight.

\footnote{HT: Text added.} Mothers who were HIV positive had infants with lower birth weight as compared to HIV. In addition, exclusive breastfeeding highly influenced child malnutrition. There was a decrease in wasting, stunting and underweight for infants who were exclusively breastfed.
Our study has shown that there was no difference in practice of exclusive breastfeeding between mothers who knew their positive HIV status and those who were HIV-uninfected for both datasets. In addition ART did not mediate the effect of knowledge of HIV status on practice of exclusive breastfeeding in 2010. The result shows that whether a mother is HIV-infected or not, it does not impact on their EBF practice. This could be as a result of impact of interventions that have raise awareness on the importance of EBF and hence mothers practice it regardless of their HIV status. Hazenba et al., in their exploratory qualitative study on promotion of exclusive breastfeeding among HIV-infected mothers in Zambia, found that gaps in understanding benefits of exclusive breastfeeding and potential for behaviour change was a challenge. They attribute lower practice of EBF among HIV-infected mothers to lack of understanding of instructions provided by health workers. Kafualufula et al., in their qualitative study done in Malawi showed that despite HIV mothers considering exclusive breastfeeding as an important component of a child’s well-being, HIV-infected mothers felt that exclusive breastfeeding led to maternal illness. Contrary to our findings, some studies have shown that HIV-infected mothers are more likely to practice exclusive breastfeeding for the first six months unlike HIV-uninfected mothers.

Propensity score based methods that aim to match confounding characteristics between exposure groups were used. All selected potential confounding factors were balanced between the two matched groups in our study. Balance was achieved for the matching methods, and significant causal estimates of maternal HIV on birth weight and knowledge of HIV status on exclusive breastfeeding were obtained for the 2015-16 MDHS data. Other studies have successfully used matching methods.
to balance differences between exposure groups [21, 34, 46]. Inverse probability weights were also used to create a pseudo-population that resulted in comparability between HIV status groups hence enabling estimating the causal effect of HIV on birth weight and exclusive breastfeeding.⁸

Alternatively, under certain conditions of exposure and risk factor association, valid causal effect estimation on health outcomes can be performed on prevalence outcomes. Assumptions needed to validate estimation of prevalence causal effects as defined by Flanders et al, [36] were defined and verified, hence enabling us to obtain prevalence odds ratios and interpret them as causal. Flanders et al, [36] states that exchangeability is an important assumption for estimation of prevalence difference in cross sectional studies, they further continue to say if confounding is present, then the exchangeability condition is not expected to hold. If one assumption is not met then the estimated causal effect could be biased [36].¹⁰ We satisfied the exchangeability assumption by controlling for observed confounders using PS matching and balance of the confounders between HIV-infected and HIV-uninfected was achieved.

We further conducted a sensitivity analysis to assess whether the causal estimates were affected by bias caused by unmeasured confounders. The sensitivity analysis showed that the p-value, for the causal estimate of birth weight between HIV infected and HIV uninfected mothers with similar 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 difference, would cause a change in the inference. This means that, though the effect size were explained by mothers HIV status, the estimates were sensitive to possible hidden bias due to unmeasured confounding.

Strengths and Limitations

One of the strengths of our study is that we used data originating from a large survey and hence a rich data with a vast number of socio-demographic characteristics of participants. The study used the propensity score matching to estimate the causal effects which can reduce confounding as observed covariates are balanced at each particular value of the propensity score [21, 29]. Several limitations of the study must be addressed.¹¹ A major weakness could be that important confounder risk factors that have been shown to be linked to birth weight were not available in this study hence affecting how the obtained estimates are sensitive to unmeasured confounders. These include maternal height, which has been shown to be associated with fetal growth measures, where shorter mothers tend to deliver with lower birth weight [47, 48].¹² Other factors include unfavourable genetic factors which predispose infants to be smaller in size and length [49]. Factors such as maternal

⁸HT: There were slight differences in causal effect estimates after matching. The IPW estimates were slightly similar to estimates for matched and unmatched sample, but the confidence intervals for the IPW were much narrower than that of the NN matching. PS matching and unmatched

¹⁰HT: Text added.
¹¹HT: Text added.
¹²HT: Text added.
nutrition, maternal passive smoking, drinking, folic acid supplement use, taking drugs, family history of birth defects are also unmeasured confounders as they are known to affect birth weight [33] and practice of exclusive breastfeeding [50]. [c5] In addition, knowledge of ART timing of initiation of therapy and type of ART regimen was not available in the data. These are known to have influenced risk of adverse birth outcomes in HIV-infected mothers. [c6] On the other hand, knowledge of date of HIV acquisition was not available in the data hence it may be possible that a mother was infected with HIV after infant birth. In this case, exposure HIV status would not affect birth weight or exclusive breastfeeding. [c7] The sample was reduced to a small size as we only looked at children who were alive and were less than 6 months of age and the data had a lot of missing values that we did not drop as we would have lost information in important variables. This might affect the robustness of the estimates obtained. The assumptions that enabled us to estimate the prevalence causal effect were not tested thus violation of the assumptions would not allow us to draw a causal conclusion.

Conclusion
Our findings suggest that [c1] using appropriate statistical tools that control for confounding in observational data, maternal HIV had a causal impact on birth weight with children born to HIV-infected mothers having higher birth weight than children born to HIV-uninfected mothers for the 2015-16 MDHS. [c2] Knowledge of mothers HIV status had an impact on practice of exclusive breastfeeding for the 2015-16 MDHS. ART had no mediating effect on the causal effect of maternal HIV infection on birth weight and on the causal effect of knowledge of HIV status on exclusive breastfeeding for the 2010 MDHS data. [c3] The study has highlighted [c4] [c5] recent success of policies and interventions such as antenatal care services that ensure HIV-infected mothers receive appropriate nutritional supplements, iron supplements and folic acid that help with their nutritional and health status and improve an unborn child’s growth. Interventions that promote practice of exclusive breastfeeding, such as counselling on infant feeding practices and health education on EBF, for the first six months among HIV-infected mothers on ART should be emphasized. This would help the child to receive appropriate nutrients and prevent illnesses such as acute respiratory infection and diarrhoea. It is important to emphasize the importance of uptake of PMTCT guidelines for effective child and maternal health outcomes. Propensity score methods act as useful tools to address the issue of effect of maternal HIV on pregnancy outcomes [c6] on survey data. Future studies therefore can look at causal effects of HIV status on pregnancy outcomes in observational studies particularly longitudinal studies and consider

[c5] HT: Text added.
[c6] HT: Text added.
[c7] HT: Text added.
[c1] HT: Text added.
[c2] HT: Text added.
[c3] HT: Text added.
[c4] HT: Text added.
[c5] HT: Text added.
methods that would handle the repeated measurement of the exposure and how to balance confounders that change with time and the same time controlling for dependence between clusters. Future Studies can also look at statistical methods such as visual diagrams, random forests to study causal effect of maternal HIV on pregnancy outcomes.

List of abbreviations
WHO - World Health Organisation
HIV - Human Immune-deficiency Virus
PMTCT - Prevention of Mother to Child Transmission
MTCT - Mother to Child Transmission
ART - Anti-retroviral Treatment
ANC - Antenatal Clinics
PS - Propensity Scores
NN - Nearest Neighbour
LMIC - Low-middle income countries

Declaration
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 [40, 51], 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

Availability of data and materials
The data used in this study can be requested and downloaded from
http://dhsprogram.com/data/dataset/Malawi_Standard-DHS_2010
http://dhsprogram.com/data/dataset/Malawi_Standard-DHS_2015

Competing interests
The authors declare that they have no competing interests.

Funding
Not applicable

Author’s 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

Acknowledgements
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.

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 Department of Statistics, University of Pennsylvania, Pennsylvania, USA. 4 Department of Statistics, University of Pretoria, Pretoria, South Africa.
30. Gu, X.S., Rosenbaum, P.R.: Comparison of multivariate matching methods: Structures, distances, and
31. Rosenbaum, P.R., Rubin, D.B.: Reducing bias in observational studies using subclassification on the propensity score. Journal of the American statistical Association 79(387), 516–524 (1984)

32. Rubin, D.B., Thomas, N.: Matching using estimated propensity scores: relating theory to practice. Biometrics, 249–264 (1996)

33. 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. Scientific reports 8(1), 5154 (2018)

34. McCaffrey, D.F., Ridgeway, G., Morral, A.R.: Propensity Score Estimation With Boosted Regression for Evaluating Causal Effects in Observational Studies. Psychological Methods 9(4), 403–425 (2004).

doi:10.1037/1082-989X.9.4.403

35. Austin, P.C.: Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statistics in medicine 28(25), 3083–3107 (2009)

36. Flanders, W.D., Klein, M., Mirabella, M.C.: Conditions for valid estimation of causal effects on prevalence in cross-sectional and other studies. Annals of epidemiology 26(6), 389–3942 (2016).

doi:10.1016/j.annepidem.2016.04.010

37. Rosenbaum, P.R.: Observational studies. In: Observational Studies, pp. 1–17. Springer, ??? (2002)

38. Rosenbaum, P.R.: Sensitivity analysis in observational studies (2005)

39. Keele, L.: An overview of rbounds: An r package for rosenbaum bounds sensitivity analysis with matched data (2010)

40. MDHS, R.: Malawi demographic and health survey 2010 report. Technical report, National Statistics Office Malawi (2010)

41. (2015), M.N.A.C.: Malawi aids response progress report 2015 (2015)

42. Corbett, M.: Support for plwhas in malawi. field exchange 25 25, 31 (may 2005)

43. Hazemba, A.N., Ncama, B.P., Sithole, S.L.: Promotion of exclusive breastfeeding among hiv-positive mothers: an exploratory qualitative study. International breastfeeding journal 11(1), 9 (2016)

44. Kafalufula, U.K., Hutchinson, M.K., Gennaro, S., Guttmacher, S.: Maternal and health care workers’ perceptions of the effects of exclusive breastfeeding by hiv positive mothers on maternal and infant health in blantyre, malawi. BMC pregnancy and childbirth 14(1), 247 (2014)

45. Oiye, S., Mwanda, W., Mugambi, M., Fitteau, S., Owino, V.: Exclusive breastfeeding is more common among hiv-infected than hiv-uninfected kenyan mothers at 6 weeks and 6 months postpartum. Breastfeeding medicine 12(5), 283–289 (2017)

46. Lanza, S.T., Moore, J.E., Butera, N.M.: Drawing causal inferences using propensity scores: A practical guide for community psychologists. American journal of community psychology 52(3-4), 380–392 (2013)

47. Inoue, S., Naruse, H., Yorifuji, T., Kato, T., Murakoshi, T., Subramanian, S., et al.: Association between short maternal height and low birth weight: A hospital-based study in japan. Journal of Korean medical science 31(3), 353–359 (2016)

48. Azcorra, H., Mendez, N.: The influence of maternal height on offspring’s birth weight in merida, mexico. American Journal of Human Biology 30(6), 23162 (2018)

49. Dubois, L., Kyvik, K.O., Girard, M., Tatone-Tokuda, F., Peruus, D., Hjelmborg, J., Skytte, A., Rasmussen, F., Wright, M.J., 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 7(2), 30153 (2012)

50. Kuchenbecker, J., Jordan, I., Rehnott, A., Herrmann, J., Jeremias, T., Kennedy, G., Muchhoff, E., Mitmuni, B., Krawinkel, M.B.: Exclusive breastfeeding and its effect on growth of Malawian infants: results from a cross-sectional study. Paediatrics and international child health 35(1), 14–23 (2015).

doi:10.1179/2046905514Y.0000000134

51. MDHS, R.: Malawi demographic and health survey 2014-15 report. Technical report, National Statistics Office Malawi (2014-15)
### Table A1: 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) | p-value | HIV-infected( %) | HIV-uninfected(%) | p-value |
|-----------------|---------------------------|--------------------------|---------|------------------|------------------|---------|
| Overall         | 58(50.0)                  | 58(50.0)                 |         | 121(25.9)        | 346(74.1)        |         |
| Residence       |                           |                          |         |                  |                  |         |
| Urban           | 15(55.6)                  | 12(44.4)                 |         | 40(26)           | 114(74)          |         |
| Rural           | 43(48.3)                  | 46(51.7)                 | 0.510   | 81(25.9)         | 232(74.1)        | 0.982   |
| Region          |                           |                          |         |                  |                  |         |
| Northern        | 9(56.3)                   | 7(43.8)                  |         | 23(29.1)         | 56(70.9)         |         |
| Central         | 19(52.8)                  | 17(47.2)                 |         | 40(24.4)         | 124(75.6)        |         |
| Southern        | 30(46.9)                  | 34(53.1)                 | 0.737   | 58(25.9)         | 166(74.1)        | 0.734   |
| Age             |                           |                          |         |                  |                  |         |
| 15-24           | 12(48.0)                  | 13(52.0)                 |         | 34(25.2)         | 101(74.8)        |         |
| 25-34           | 30(50.9)                  | 29(49.1)                 |         | 55(25.6)         | 160(74.4)        |         |
| 35 above        | 16(50.0)                  | 16(50.0)                 | 0.972   | 32(27.4)         | 85(72.7)         | 0.915   |
| Wealth          |                           |                          |         |                  |                  |         |
| Poor            | 13(44.8)                  | 16(55.2)                 |         | 35(24.8)         | 106(75.2)        |         |
| Middle          | 12(52.2)                  | 11(47.8)                 |         | 24(30.4)         | 55(69.6)         |         |
| Rich            | 33(51.6)                  | 31(48.4)                 | 0.812   | 62(25.9)         | 185(74.9)        | 0.609   |
| Education       |                           |                          |         |                  |                  |         |
| None            | 7(41.2)                   | 10(58.8)                 |         | 20(24.1)         | 63(75.9)         |         |
| Primary         | 39(49.4)                  | 40(50.6)                 | 0.511   | 62(27.7)         | 162(72.3)        |         |
| Secondary       | 12(60.0)                  | 8(40.0)                  |         | 38(26.4)         | 106(73.6)        | 0.290   |
| Anaemia         |                           |                          |         |                  |                  |         |
| No              | 29(47.5)                  | 32(52.5)                 |         | 61(24.2)         | 191(75.8)        |         |
| Yes             | 29(52.7)                  | 26(47.3)                 | 0.577   | 60(27.9)         | 155(72.1)        | 0.363   |
| Had fever last 2 weeks |               |                          |         |                  |                  |         |
| No              | 27(48.2)                  | 29(51.8)                 |         | 72(96.2)         | 203(73.8)        |         |
| Yes             | 31(51.7)                  | 29(48.3)                 | 0.710   | 48(25.4)         | 141(74.6)        | 0.941   |
| Had Diarrhoea   |                           |                          |         |                  |                  |         |
| No              | 43(49.4)                  | 44(50.6)                 | 0.830   | 44(27.9)         | 114(72.1)        | 0.749   |
| Yes             | 15(51.7)                  | 14(48.3)                 |         |                  |                  |         |

### Table A2: Comparison of Covariates for HIV status before and after Propensity Score Stratification

| Covariate  | p-value(2010)) | p-value(2015-16) |
|------------|----------------|-----------------|
|            | [c3]Before stratification | [c4]After stratification |
| Age        |               |                 |
| 15-24      | [c5]<0.001    | [c6]<0.001      |
| 25-34      | [c7]<0.001    | [c8]<0.001      |
| 35 above   | [c9]<0.001    | [c10]<0.001     |
|            | [c11]<0.001   | [c12]<0.001     |
| Wealth     |               |                 |
| Poor       | [c13]<0.553   | [c14]<0.440     |
| Medium     | [c15]<0.02    | [c16]<0.817     |
| Rich       | [c17]<0.01    | [c18]<0.618     |
| Residence  |               |                 |
| Urban      | <0.001        | 0.424           |
| Rural      | <0.001        | 0.424           |
| Region     |               |                 |
| Northern   | 0.001         | 0.514           |
| Central    | 0.001         | 0.681           |
| Southern   | <0.001        | 0.368           |
| Anemia     |               |                 |
| No         | 0.04          | 0.537           |
| Yes        | 0.04          | 0.537           |
Table A3: HIV status Group Differences on Covariates age, residence, region and anemia, by Strata Based on Propensity Score Quintiles on the matched sample for 2010 MDHS

| Distribution | Sample (N) | P-value |  |  |  |  |
|--------------|------------|---------|---|---|---|---|
|              |            | HIV-uninfected | HIV-infected | age | Residence | Region | Anaemia |
| Quintile 1   | 12         | 12       | 1.00 | 1.00 | 1.00 | 1.00 |
| Quintile 2   | 12         | 12       | 0.497| 0.653| 0.513| 0.653|
| Quintile 3   | 11         | 11       | 0.895| 0.392| 0.670| 1.00 |
| Quintile 4   | 11         | 12       | 0.901| 0.354| 0.924| 0.901|
| Quintile 5   | 12         | 11       | 0.538| 0.904| 0.484| 0.408|

Table A4: HIV status Group Differences on Covariates age, residence, region and anemia, by Strata Based on Propensity Score Quintiles on the matched sample for 2015-16 MDHS

| Distribution | Sample (N) | P-value |  |  |  |  |
|--------------|------------|---------|---|---|---|---|
|              |            | HIV-uninfected | HIV-infected | age | Residence | Region | Anaemia |
| Quintile 1   | 72         | 24       | 0.666| 1.00 | 0.761| 0.604|
| Quintile 2   | 69         | 23       | 0.937| 1.00 | 0.885| 0.889|
| Quintile 3   | 72         | 24       | 0.87 | 0.604| 0.852| 1.00 |
| Quintile 4   | 66         | 24       | 0.785| 0.868| 0.227| 0.613|
| Quintile 5   | 67         | 26       | 0.806| 0.903| 0.394| 0.526|