Methodological considerations related to the epidemiologic study of birth outcomes: maternal ART use and adverse birth outcomes

Thokozile Rosemary Malaba

MLBTHO002

Thesis presented for the degree of

Doctor of Philosophy

Division of Epidemiology and Biostatistics

School of Public Health and Family Medicine

Faculty of Health Sciences, University of Cape Town

Date: 5th May 2023

Supervisors:

Professor Landon Myer

Professor Marie-Louise Newell

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

PREAMBLE

Abstract

Background

A major factor contributing to continued high under-five mortality in sub-Saharan Africa (SSA), is maternal HIV infection, which is associated with adverse birth outcomes such as preterm delivery (PTD), small-for-gestational age (SGA) and low birthweight (LBW) infants. Introduction of antiretroviral therapy (ART) during pregnancy has been a successful intervention for promotion of maternal and infant health, however it has also been linked to an increased risk of adverse birth outcomes. Consequently, the association between maternal ART use and these adverse outcomes is an important area of research in SSA which has the majority of pregnant women living with HIV and the highest rates of PTD, SGA and LBW infants. However, the current state of epidemiologic knowledge remains limited because most evidence on this association comes from observational studies, which have previously given inconsistent findings. Accordingly, the overarching aim of this PhD was to reliably quantify the relationship between maternal ART use and adverse birth outcomes, by addressing the role of methodological factors inherent in observational research in this association.

Methods

This research included pregnant women (aged ≥18 years) seeking antenatal care at a public sector midwife obstetric unit in Gugulethu (GMOU), Cape Town, enrolled into two separate dedicated research cohort studies. Enrolled women were followed-up during pregnancy and postpartum with their infants, with data obtained from study questionnaires, physical examinations and abstraction of clinical and obstetric records. In parallel, routine electronic data, linked across clinics and data sources were obtained for all pregnant women at the GMOU (pregnancy exposure registry (PER)) and across the Western Cape province (PHDC).

Findings

The incidence of gestational age (GA) based birth outcomes and the association between maternal ART use and these outcomes, was found to be substantially influenced by method of GA assessment used. While GA based on both last menstrual period (LMP) and measurement of symphysis fundal height (SFH) led to under and over estimation (relative to ultrasound), only LMP-based GA gave rise to a biased measure of association.

Across data sources used in this research, an overall PTD incidence of 17% was observed which was lower than incidences observed in the pre-universal ART era. In the cohort studies, there was no

significantly increased PTD risk in women living with HIV (compared to living without HIV), predominantly on the tenofovir + emtricitabine + efavirenz regimen. However, when assessed in the significantly larger population of pregnant women (PHDC), an increased PTD risk in women living with HIV was observed. There did not appear to be differences in PTD risk by ART status in the cohort studies or PER. However, across the province those initiating ART preconception were at increased PTD risk compared to those initiating during pregnancy. Blood pressure, particularly when assessed longitudinally played an important role in the association between maternal ART use and PTD, high normal and abnormal blood pressure trajectories associated with increased PTD risk. There did not appear to be any effect modification in the trajectory groups by HIV status for PTD.

In the cohort study the overall incidence of SGA infants was 9%, with an increased SGA risk observed in women living with HIV compared to living without HIV. While no differences were observed in SGA risk by ART status, the highest risk was observed among women initiating in the second trimester. An overall LBW incidence of 13% was observed in the cohort study, with no differences observed in risk by HIV status or ART status. Blood pressure also played a role in the occurrence of LBW infants, with abnormal trajectories associated with increased LBW infant risk. Additionally, effect modification among women with abnormal trajectory groups, those living without HIV were at increased risk of LBW infants compared to those living with HIV.

Finally, investigating this association using both cohort studies and population based electronic health care data proved to be valuable. The three data sources gave similar effect estimates, with varying levels of precision, and each with distinct but complementary benefits. The cohort studies and PER included smaller select groups of women and provided detailed investigation of risk factors that could impact the overall association. In contrast the provincial dataset, had limited risk factor data, but provided overall associations with the ability to detect subtle differences.

Conclusion

Reassuringly, the magnitude of difference in PTD risk by HIV status under policies of universal maternal ART use, appears to be smaller than in the past. Of concern, however, was the finding of increased SGA risk. Taken together, these findings highlight the need to improve mechanistic understanding of ART-mediated adverse birth outcomes, in order to ensure optimal maternal and infant outcomes. The methodological findings underscore the importance of considering the potential for bias related to selection and measurement when designing and/or evaluating findings from studies investigating this association and by extension other medications in pregnancy.

Dedication

To my mother **Lucie Carolyn Malaba**

For your unconditional love and always inspiring me to be the best I can be.

and to my beloved nieces, **Ayanda and Zandile Malaba**

a reminder that there is no limit to what we, as women, can accomplish.

Acknowledgement

This thesis would never have been possible, and I would not be where I am without the contributions, support, and encouragement of so many others.

Thank you to my supervisors, for sharing your knowledge and your constant commitment to my development. Landon, my sincerest gratitude for your mentorship and continued guidance, and most importantly for directing me onto the road that has led me to this point. You saw me where I am now, long before I saw it as a possibility, and I am privileged to have been your student. Marie-Louise, you expected the best from me and guided me while also encouraging me to go my own way. Your kind words of support have been truly invaluable.

To the PIMS and B-Positive study teams, notable Hlengiwe, Colin, Megan, Lee-Ann, Emma and Dineo, working with each one of you has taught me so much and I am truly grateful. Thank you to the clinic staff at the Gugulethu Midwife Obstetric Unit who have been so supportive of our work over the years. I would also like to acknowledge all the women who took part in the studies that provided data for this thesis, I dedicate this work to you.

Thank you to my friends and family, many of whom have been neglected lately, for your support in so many ways and for keeping me grounded in everything that wasn't PhD-related. To my siblings Tshala, Mjele and Praise, and my extended family in Zimbabwe, South Africa and the United Kingdom. In particular, Mduduzi and Zandile, Nothando, Josephine, and my aunts Agnes, Maria, Margaret, Jacqueline and Joyce. Special mention to Nomkhosi for putting up with my late night calls. To all my amazing friends, in particular my sounding boards Cynthia, Jasantha, Vundli and Sandra. Also, Wadzanai, Dadirai, Daphine, Priscilla, Kundai, Donald, Chido and Hajierah. To those I have not mentioned by name, I also wish to express my deepest gratitude.

Lastly my heartfelt appreciation goes to my parents Joseph and Lucie. You have both been my source of inspiration, support, and guidance. You have taught me to be unique, determined, to believe in myself, and to always persevere. I am truly thankful and honoured to have you as my parents.

"To maximise the benefit to society, you need to not just do research, but do it well".

Doug Altman

Preface

This thesis is presented in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) in the Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town. The work included in this thesis is original research, and has not, in whole or in part, been submitted for another degree at this or any other university. The contents of this thesis are entirely the work of the candidate or, in the case of multi-authored published papers, constitutes work for which the candidate was the lead author.

This thesis includes published manuscripts, as per general provision 6.7 in the General Rules for the Degree of Doctor of Philosophy (PhD) of the University of Cape Town. I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publications in my PhD thesis, and where co-authorships are involved, my co-authors have agreed that I may include these publications. The following manuscripts (four published, one to be submitted) are included in the thesis, and are presented as self-contained chapters in the following order:

1. **Malaba TR**, Newell ML, Myer L, Ramokolo V. Methodological considerations for preterm birth research. *Front Glob Womens Health*. 2022;2:821064.
2. **Malaba TR**, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol.* 2018; 28(12):893-900.
3. **Malaba TR**, Mukonda E, Matjila M, Madlala HP, Myer L, Newell ML, PIMS Study Group. Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age. *Paediatr Perinat Epidemiol*. 2022;36(4):525-535.
4. **Malaba TR,** Cois A, Madlala HP, Matjila M, Myer L, Newell ML. Blood pressure trajectories during pregnancy and associations with adverse birth outcomes among HIV-infected and HIV-uninfected women in South Africa: a group-based trajectory modelling approach. *BMC Pregnancy Childbirth*. 2020;20(1):742-754.
5. **Malaba TR,** Kalk E, Madlala HP, Davies MA, Boulle A, Myer L. Impact of data sources and selection bias on the association between maternal ART use and preterm delivery. *Manuscript to be submitted in 2022.*

In addition, the following published manuscript is directly related to the thesis and included as an appendix:

6. **Malaba TR**, Myer L, Gray CM, Newell ML. Cohort profile: Prematurity Immunology in Mothers living with HIV and their infants Study (PIMS). *BMJ Open.* 2021;11(9):e047133.

The contribution of the candidate to each manuscript is outlined at the start of each chapter. The candidate was the lead and corresponding author on all manuscripts, prepared the datasets for analysis, conducted all analyses except for the multiple overimputation in Chapter 5 (which was conducted together with Elton Mukonda), and drafted all versions of the manuscripts. All co-authors reviewed and approved the submitted manuscripts and the candidate reviewed co-author comments and integrated them into the manuscripts prior to submission.

The data for analyses in this thesis originated from data collect from two studies which were conducted consecutively: Prematurity Immunology in Mothers living with HIV and their infants Study (PIMS), and A population-based evaluation of expanded (antiretroviral therapy) ART access in pregnancy (B positive). For the PIMS study, the candidate was employed as the study coordinator on this project from 2014 to 2018 and played a central role in data collection, study implementation and data analysis. For the B-Positive study, the candidate was played a key role in designing data collection tools and was involved with data collection and analysis. Throughout the course of these studies, she developed, in conjunction with her supervisors, the independent concepts for the analyses presented in this PhD thesis. The candidate's primary supervisor (Prof Myer) has confirmed to the University of Cape Town Doctoral Degrees Board that the included manuscripts all overwhelmingly reflect the candidates own scientific work.

Student Name: Thokozile Rosemary Malaba

Date: 5th May 2023

Signed:

Signed by candidate

Per University guidelines, the text of each publication is presented verbatim in this thesis. As such, there are minor discrepancies in terminology between papers; these have not been changed in order to reflect the publications. Minor changes have been made to style and to figure and table numbers to ensure consistency throughout this thesis.

List of abbreviations

AC	Abdominal circumference
AGA	Appropriate for gestation age
ANC	Antenatal care
AOR	Adjusted odds ratio
APGAR	Appearance, pulse, grimace, activity and respiration
ART	Antiretroviral Therapy (triple drug)
ARR	Adjusted risk ratio
ARV	Antiretroviral
AUDIT-C	Alcohol use disorders identification test-concise
BIC	Bayesian information criterion
BP	Blood pressure
BPD	Biparietal diameter
BMI	Body mass index
CD4	Cluster of differentiation 4
CI	Confidence interval
CRL	Crown rump length
CRVS	Civil registration and vital statistics
DAG	Directed acyclic graph
DBP	Diastolic blood pressure
DOH	Department of Health
DTG	Dolutegravir
EFV	Efavirenz
EFW	Estimated fetal weight
HER	Electronic health records
ELBW	Extremely low birth weight
ELISA	Enzyme-linked immunosorbent assay
EPDS	Edinburgh postnatal depression scale
EPL	Early pregnancy loss
FL	Femur length
FTC	Emtricitabine
GA	Gestational age
GBTM	Group-based trajectory modelling

GMOU	Gugulethu midwife obstetric unit
GWG	Gestational weight gain
HBW	High birthweight
HC	Head circumference
HCG	Human chorionic gonadotrophin
HDP	Hypertensive disorders in pregnancy
HIC	High income country
HIV	Human Immunodeficiency Virus
INSTI	Integrase strand transfer inhibitors
IPW	Inverse probability weighting
IUGR	Intrauterine growth restriction
LBW	Low birthweight
LGA	Large for gestational age
LMIC	Low to middle income countries
LMP	Last menstrual period
LPV/r	Lopinavir/Ritonavir
MAP	Mean arterial pressure
MCR	Maternity case record
MDG	Millennium development goal
ML	Machine learning
MOU	Midwife obstetric unit
MVM	Maternal vascular malperfusion
NMR	Neonatal mortality rate
NPV	Negative predictive value
NTD	Neural tube defects
NRTI	Nucleoside/nucleotide reverse transcriptase inhibitors
NNRTI	Non-nucleoside reverse transcriptase inhibitors
NVP	Nevirapine
OR	Odds ratio
PER	Pregnancy exposure registry
PHDC	Provincial health data centre
PI	Protease inhibitors
PMTCT	Prevention of mother to child transmission
PTD	Preterm delivery

PPV	Positive predictive value
PrEP	Pre-exposure prophylaxis
PUL	Pregnancy of unknown location
RCT	Randomised control trial
RR	Risk ratio
SBP	Systolic blood pressure
SE	Sensitivity
SES	Socio-economic status
SDG	Sustainable development goals
SFH	Symphysis-fundal height
SGA	Small for gestational age
SSA	Sub-Saharan Africa
SP	Specificity
TAF	Tenofovir alafenamide
TB	Tuberculosis
TDF	Tenofovir disoproxil fumarate
UNAIDS	Joint United Nations Programme on HIV/AIDS
US	Ultrasound
U5MR	Under-five mortality rate
VL	Viral load
VLBW	Very Low birthweight
WC	Western Cape
WHO	World Health Organization
WLHIV	Women living with HIV
ZDV	Zidovudine
3TC	Lamivudine

Table of contents

List of tables

List of figures

List of supplementary tables

List of supplementary files

Chapter 1: Introduction

1.1 Introduction

1.1.1 Overview

Globally, there were 1.3 million pregnant women with HIV in 2020, of which the overwhelming majority reside in sub-Saharan Africa (SSA) (1). Provision of antiretroviral therapy (ART) as part of the continuum of care in pregnancy has been one of the most important interventions implemented in response to the HIV epidemic, as part of prevention of mother to child transmission programmes. The 2013 WHO treatment guidelines which recommended that all pregnant HIV-infected women receive ART for life (Option B+) (2), led to significantly increased proportions of pregnant women receiving ART. Consequently in 2020, 85% of HIV-infected pregnant women had access to ART, which has contributed to significantly improved maternal health and substantial decreases in perinatal transmission of HIV (3). However, pregnancy presents a unique situation because both maternal and fetal factors require consideration, so while the benefits of ART are indisputable, significant concerns remain around fetal exposure to HIV and antiretroviral drugs. Of particular concern is that maternal ART use has been linked with an increased risk of adverse birth outcomes which has far-reaching consequences as these outcomes are major contributors of neonatal and child mortality and morbidity. Consequently, at a population level, any excess risks of these adverse outcomes is worrying because it is likely to substantially impact the success of meeting the Sustainable Developmental Goals (SDG) in sub-Saharan Africa generally and in South Africa specifically.

This chapter introduces the PhD thesis, and covers a broad overview of child and neonatal mortality, a major public health concern particularly in low-to-middle income countries (LMIC). This chapter will also cover the main exposures and outcomes of the thesis, which are all major contributory factors for the continued high under-five mortality rate (U5MR) in SSA. The main outcomes of interest in this thesis are adverse birth outcomes, namely preterm delivery (PTD), low birth weight (LBW) and small for gestation age (SGA); which account for significant proportion of neonatal deaths occurring in the first weeks of life. The main exposure of interest in this thesis is maternal HIV infection and ART use, as maternal infections and their treatment play a substantial role in the occurrence of adverse birth outcomes particularly in South Africa where there is a high prevalence of HIV. Using the association between maternal ART use and adverse birth outcomes as an example, this thesis will touch on several methodological issues that may arise when investigating the association between similar exposures and birth outcomes. The aims and objectives of this thesis,

linked to the specific methodological issues, will be highlighted in this chapter together with a description of the data sources that will be used. Lastly this chapter will also set out the papers that form the content of this PhD thesis.

1.1.2 Progress of efforts to reduce neonatal and child mortality in sub-Saharan Africa

Since 1990 substantial global progress has been made in reducing childhood mortality, with the U5MR declining 61% from 93.0 (91.7-94.5) to 37.7 (36.1-40.8) deaths per 1000 livebirths between 1990 and 2020 (4-6) (Figure 1.1, top panel). This decline is a reflection of the efforts across many platforms, by a range of organisations and governments, in the Millennium Development Goals (MDG) era (2000–2015),-and the subsequent SDG era (2016 onwards) (4). Sub-Saharan Africa continues to have the highest rates of U5MR, estimated at 75.8 (70.2–85.9) deaths per 1000 livebirths in 2019 (5). Of concern is that SSA (55%) together with South Asia (26%) accounted for the majority of global deaths despite only accounting for half of the global U5 population (6). However, encouragingly, SSA has seen declines of 59% from 1990, following the same pattern of decline globally: fastest pace of decline between 2000 and 2009, followed by a possible slowdown between 2010 and 2019 (5). With declines in the level of U5 mortality, a larger proportion of deaths are concentrated in the neonatal period, the riskiest period for child survival, in 2019 (47%). This is an increase from 1990 (40%) and means the neonatal (<28 days) decline is being outpaced by declines infancy and childhood (1 – 59 months) (5).

Between 1990 and 2019 the global neonatal mortality rate (NMR) declined by 53% from 36.6 (35.6-37.8) to 17.5 (16.6-19.0) deaths per 1000 live births (5, 7) (Figure 1.1, bottom panel). While this NMR decline was not as quick as that for U5MR, the global pattern in decline in NMR followed that of U5MR, with a quickening in the pace of decline in the 2000s and possible slowdown in the most recent decade (5). As with U5MR, neonatal mortality is regionally concentrated, with SSA having the highest global NMR in 2019, at 27.5 (25.2–31.7) deaths per 1000 livebirths, a 41% reduction since 1990 (5, 6).

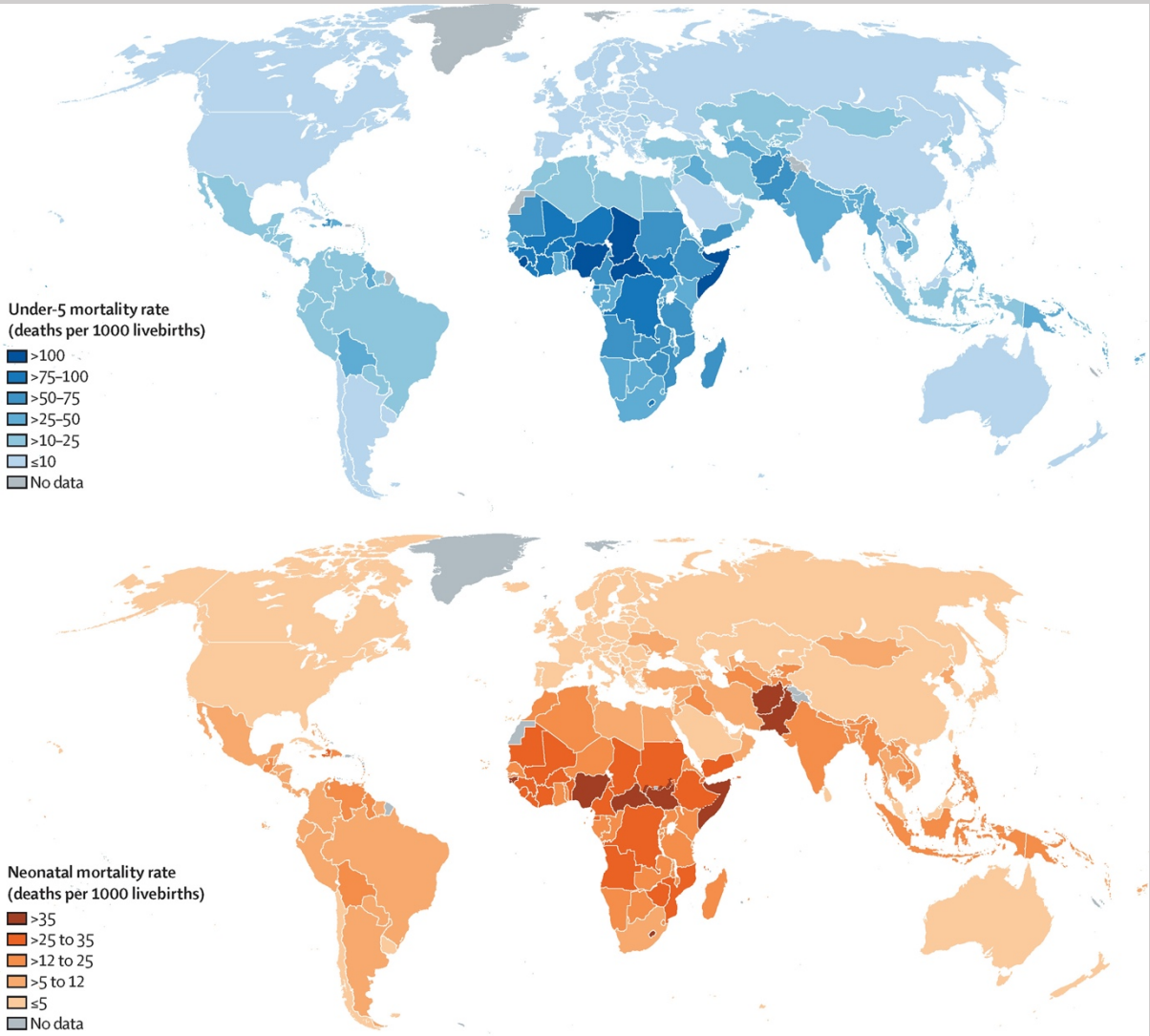

Figure 1.1 Global under-5 mortality rate (top panel) and neonatal mortality rate (bottom panel) in 2019

From: Sharrow D, Hug L, You D, et al. Global, regional, and national trends in under-5 mortality between 1990 and 2019 with scenario-based projections until 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *Lancet Glob Health.* 2022;10(2):e195-e206.

While significant progress has been made, these declines fell short of the MDG target of a two-thirds reduction in U5MR by 2015 (8, 9); with the U5MR only declining by approximately half in the MDG period (10). In the post-MDG era, there was an understanding that continued focus on improving child health and survival would be required, particularly in SSA. This forms the basis of SDG 3 which aims to achieve good health and wellbeing for all by 2030, adopted by the United Nations in 2015 (11). The second target of SDG3 related to child survival, aims to end preventable newborns and U5 deaths, by achieving a U5MR of ≤25 deaths (SDG target 3.2.1) and an NMR of ≤12 deaths (SDG target 3.2.2) per 1000 livebirths (12, 13). Almost 75% of countries at risk of missing the child survival SDG targets are in SSA (5, 12); highlighting the need for increased attention to this early period of life and urgent action to prevent newborn deaths (6).

1.1.3 Adverse birth outcomes

1.1.3.1 Contribution of adverse birth outcomes to continued high neonatal and child mortality

Under-five deaths taking place across different ages have distinct causes requiring varied interventions (14). Consequently disaggregation of deaths by age group is essential, as seen in the SDGs where neonatal mortality reductions are explicitly listed as goals, separately from goals of U5M reductions (15). Reducing neonatal mortality is widely considered to be 'unfinished business' of the SDGs, as these deaths disproportionately occur in LMICs (16). A major factor contributing to the continued high neonatal and U5 mortality in LMICs is adverse birth outcomes in particular complications of preterm delivery (Figure 1.2), which together with other adverse birth outcomes serve as markers of newborn vulnerability, with short and long-term outcomes (7). Vulnerable infants who manage to survive the neonatal period are not in the clear as they are still at increased risk of morbidity mortality, and sub-optimal growth and development (17, 18).

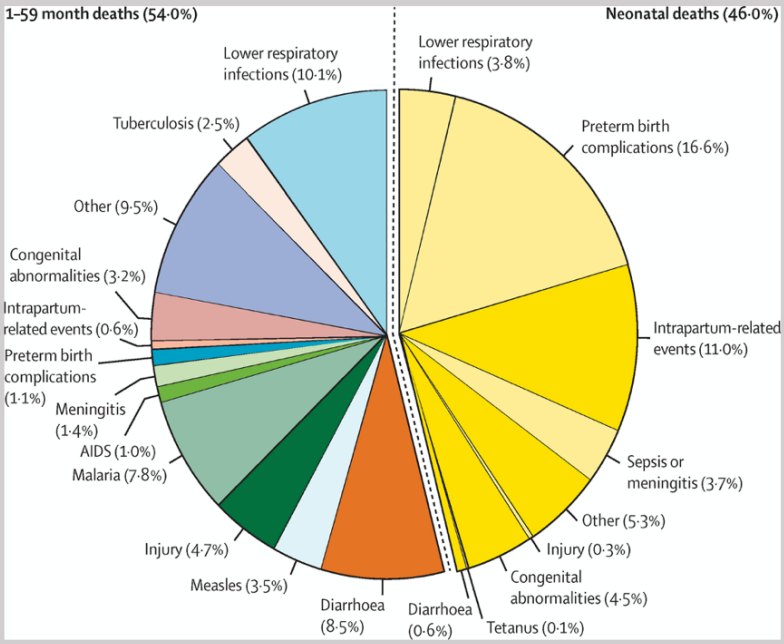

Figure 1.2 Global distribution of causes of under-5 and neonatal mortality

From: Perin J, Mulick A, Yeung D, et al. Global, regional, and national causes of under-5 mortality in 2000–19: An updated systematic analysis with implications for the Sustainable Development Goals. *Lancet Child Adolesc Health.* 2022;6(2): 106-115

Of most concern are infants who are born too early, because complications from PTD are the most important direct cause of death in the neonatal period accounting for as much as 35% of all neonatal deaths (19, 20). Complications from PTD are also is the second largest cause of under-5 deaths, accounting for 18% of all these deaths (19, 20). The other adverse outcomes which are of interest in this thesis also contribute to increased risk of neonatal mortality, with mortality rates increasing proportionally with decreasing gestational age or birthweight (21, 22). Infants that are SGA, likely due to intrauterine growth restriction (IUGR), are at a considerably high risk of neonatal morbidity and mortality because they are more likely to have morbidities which confer a higher risk of death (21). Additionally, consideration confers further attributable risk, as SGA fetuses are up to four times more likely to be stillborn (23). LBW, which can be due to IUGR and/or infants being born preterm, is an indirect cause of neonatal deaths, with 80% of neonatal deaths occurring in this group

of infants (17, 24). Infants with one or more adverse birth outcomes are at a higher risk for neonatal mortality, for example infants who are both preterm and SGA are at 15 times increased mortality risk than term and appropriate size for gestational age (AGA) infants (17, 21, 25-27). Advances in neonatal health care and collaborative efforts of obstetricians and neonatologists have improved survival rates of preterm, LBW and/or SGA infants. However, these infants remain susceptible to a variety of subsequent health and developmental problems particularly related to organ immaturity (28), including respiratory, immunologic, nervous system, and behavioural problems (26, 29).

1.1.3.2 Most common aetiological pathways of adverse birth outcomes in sub-Saharan Africa

Despite the three adverse birth outcomes of interest being linked, their aetiologies are multifactorial and incompletely understood (30). So while these outcomes may have common risk factors related to sociodemographic, obstetric and other health-related factors, their relative importance for PTD, LBW and SGA differ (31, 32). Additionally these risk factors may not be the same across different resource settings or socioeconomic statuses within a society. For example, while maternal undernutrition can be related to all three outcomes, hypertensive disorders of pregnancy and maternal infections are the immediate risk factors leading to PTD in LMICs (33).

Maternal infections play a substantial role in the occurrence of adverse birth outcomes in SSA where infections such as syphilis, HIV and malaria are widely prevalent (34-37). A growing body of evidence has linked these maternal infections with abnormal placental pathology and altered maternal and fetal haemodynamics (38). A common mechanism underlying adverse birth outcomes is inadequate placental and fetal development which can be facilitated by infection-driven dysregulation of inflammation at the materno-fetal interface (38-40) (Figure 1.3).

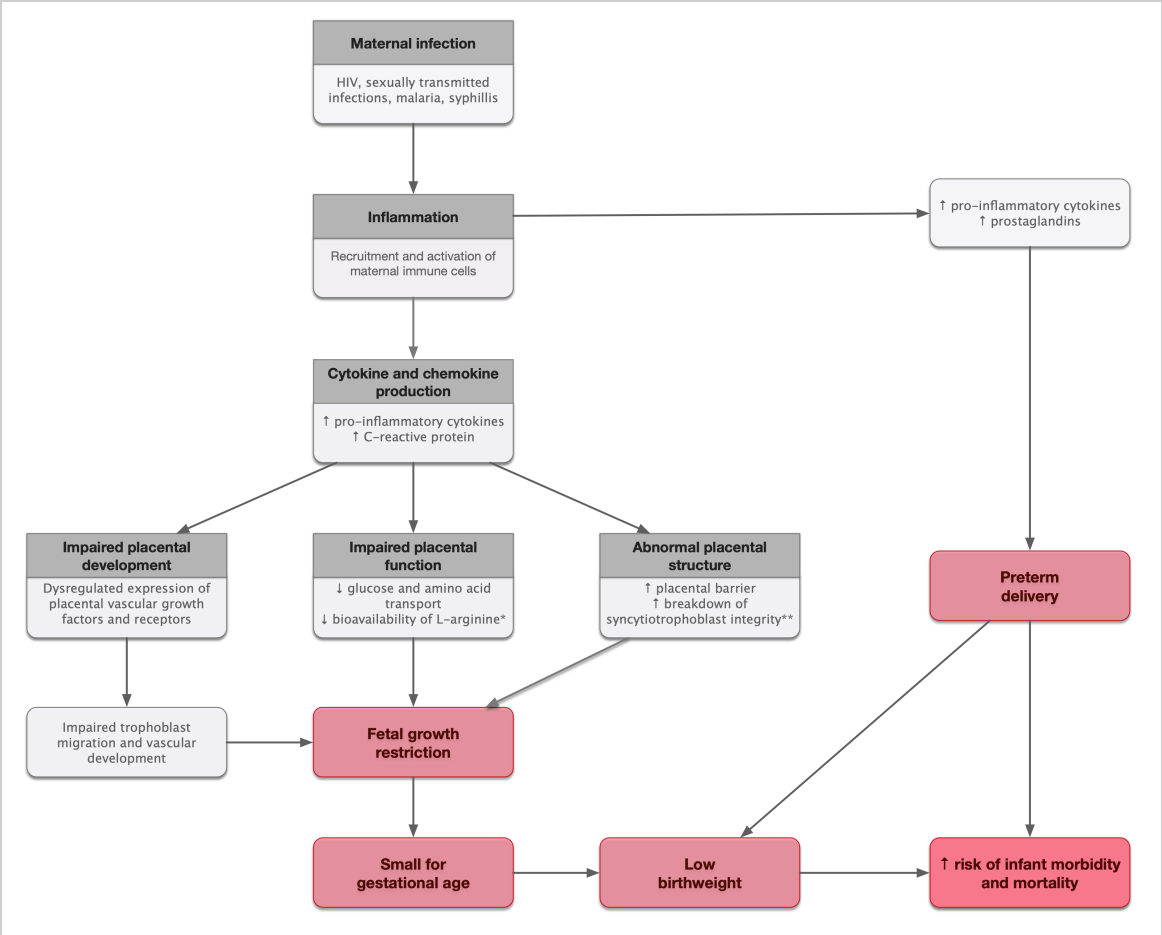

Figure 1.3 Pathways leading to infection-associated adverse birth outcomes

Inflammatory events include secretion of cytokines and chemokines shown to have been associated with adverse birth outcomes. The increased pro-inflammatory environment has been associated with: (1) dysregulated growth hormone and vascular factor expression; (2) abnormal placental development; and (3) reduced placental nutrient transport contributing to placental insufficiency.

* L-arginine promotes intrauterine growth
** Syncytiotrophoblast: placental barrier that allows nutrient exchange and plays vital roles in fetal growth

Adapted from: Chua CLL, Hasang W, Rogerson SJ, Teo A. Poor Birth Outcomes in Malaria in Pregnancy: Recent Insights Into Mechanisms and Prevention Approaches. *Front Immunol.* 2021;12:621382.

Additionally, medications used in pregnancy either to treat maternal infections or other conditions are also known to cause adverse birth outcomes through the disruption of specific steps in fetal development. While mechanisms through which maternal infections such as HIV are known, the mechanisms through which medications such as ART can increase the risks of adverse birth outcomes are less well understood (41). The global adverse birth outcome burden remains high, partly because of critical knowledge gaps that exist in our understanding of the mechanisms underlying these outcomes, and consequently limited or ineffective mitigation strategies (38). Given that there is no question ART be recommended in pregnancy, and findings that specific ART

regimens may be associated with worse adverse birth outcomes, increased efforts will be required to undertake prospective studies and birth surveillances as new ART regimens are introduced (42).

1.1.3.3 Overview of methodological challenges in studying adverse birth outcomes

Most of the evidence on the impact of exposures in pregnancy such as ART use and adverse outcomes comes from observational studies because pregnant women are typically excluded from clinical trials. This introduces complexity in attributing causality in the relationship between maternal ART use and the occurrence of adverse birth outcomes, consequently an overview of common methodological issues pertinent to the association of interest in this thesis is warranted. The main methodological challenges of interest in this thesis are those that can result in selection or measurement bias which could impact findings of the association between maternal HIV infection and ART use and adverse birth outcomes. In order to understand the predictors and antecedents of any adverse birth outcomes and related biological processes, it would be advantageous to investigate these through well designed studies consisting of representative samples of pregnancies with robust measurement of preconception and/or pregnancy exposures, covariates and outcomes of interest.

It is implicitly assumed in observational studies that the data are representative of the general population. However, the dynamic and complex nature of study populations in studies investigating pregnancy and birth outcomes, which include women who may conceive or who are pregnant, fetuses or neonates, may undermine this assumption, inducing selection bias (43). This can be through the processes of selection and attrition inherent in these populations which render them

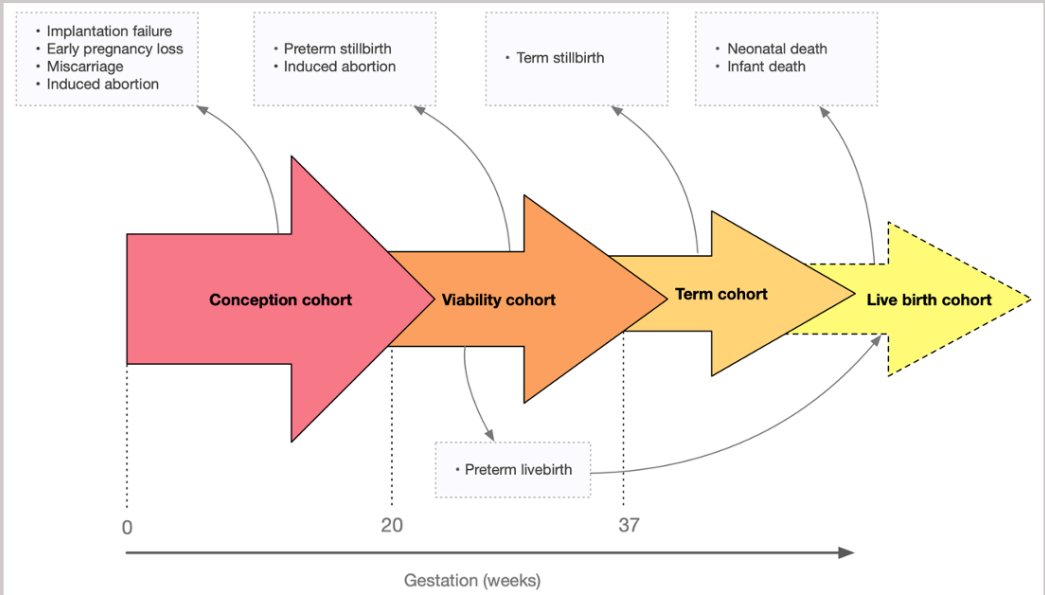

Figure 1.4 Selection and attrition processes during the course of human reproduction and gestation

Adapted From: Snowden JM, Bovbjerg ML, Dissanayake M, Basso O. The curse of the perinatal epidemiologist: inferring causation amidst selection. *Curr Epidemiol Rep*. 2018;5(4):379-387.

incompletely observable due to the extended process of attrition from embryonic development through to birth (43, 44) (Figure 1.4). Additionally, studies investigating birth outcomes are subject to further selection, because women enrolled are usually attending routine care services, meet study specific eligibility criteria, have an exposure medication of interest, or have an outcome of interest – which does not give a representative sample.

Measurement tools that are robust, valid, and standardised are critical for the avoidance of misclassification bias (45). However these are not always available for observational studies as such, interpretation of findings from the early literature on the association of maternal HIV infection and ART use and adverse birth outcomes is complicated by inconsistent definitions and measurements of both the exposure and outcomes. The primary measurement concern in this thesis relates to the imprecision surrounding the measurement of gestational age, as this impacts both exposure and outcome classification. Mismeasurement of gestational age has major implications for outcome ascertainment with misclassification likely particularly in gestation-based outcomes such as PTD and SGA. Determination of the exact timing of exposures and outcomes is another measurement concern which is also linked to gestational age mismeasurement. Well defined timing of exposure is required because the stage of fetal development at which the exposure takes place has implications for types of adverse outcomes that could occur (46). Additionally, some outcome timings such as pregnancy losses are difficult to empirically measure (particularly at population levels) and rely on proxy markers which increases the risk of outcome misclassification (47).

The unique features of pregnancy and its outcomes, as well as the complex web of interrelated factors surrounding the putative causal effects of maternal HIV/ART on adverse birth outcomes introduce some analytical complexity (48, 49). Consequently, a considered analytical approach is required, involving exploration of the associations between the exposure, outcome, and potential third variables (confounders, mediators and effect modifiers). Failure to account for the methodological and analytic challenges of pregnancy through inappropriate adjustment of variables which are either measured after the exposure has started or are potential mediators of the causal exposure-outcome relationship can introduce bias and hinder the interpretability of results and muddy actionable conclusions (50).

1.2 Problem statement and rationale

In contrast to other SSA countries, South Africa's high under-5 and neonatal mortality problem (U5 rate of 34.5 (30.5-38.8) deaths per 1000 livebirths in 2019), has occurred against a backdrop of sustained economic growth (51). As an upper middle-income country, these rates would be

expected to be considerably lower, and this incongruence can at least be attributed to the impact of the HIV and tuberculosis epidemics which have derailed progress in the reduction of child mortality. HIV directly affects infant survival through mother-to-child HIV transmission (MTCT) and increasing the risk of adverse birth outcomes, and indirectly through maternal mortality.

While the success of ART in pregnancy for improved maternal health and prevention of MTCT is indisputable, a major challenge remaining which needs addressing is the ART-associated increased risk of adverse birth outcomes as this directly impacts neonatal and child mortality rates. In recent times, there has been an increase in larger studies being conducted in SSA, the region with the highest HIV burden, to examine the association between maternal HIV infection and ART use and adverse birth outcomes. Despite this the current state of epidemiological knowledge remains limited, because of the use of predominantly observational study designs. Consequently, data from well-designed, sufficiently powered, prospective studies from SSA are still limited. Previous findings for the association between maternal ART use and adverse birth outcomes have been inconsistent across different settings, with the presence of selection and measurement bias issues linked to data sources, exposure, covariate and outcome ascertainment, and analytical choices thought to contribute to this variability.

In this treat-all era, as there is an expansion of ART options pregnant women will continue to be a priority population for the optimisation of treatment in order to promote their own health and that of their ART-exposed infants. Consequently addressing adverse outcomes linked to ART in pregnancy is of significant public health importance and will require a deeper understanding of the burden of adverse birth outcomes and their associated determinants through well designed reliable studies and analyses. This understanding could also shed light on which actions should be prioritised to enable significant reduction of child and neonatal mortality in South Africa and the SSA region at large.

1.3 Aims and objectives

The overall aim of this PhD research is to more reliably quantify the relationship between maternal ART use[1] and adverse birth outcomes (preterm delivery[2], small for gestational age and low birth weight infants), through addressing the role of methodological factors inherent in observational

[1] In this thesis all women living with HIV unless otherwise stated are treated with ART.

[2] Throughout the thesis, the outcome preterm delivery (event taken from the maternal point of view) is predominately used, except for Chapter 3 where preterm birth (event taken from the infant point of view) is used.

research of birth outcomes in general and specifically for the association of interest. In particular, how measurement and selection related issues impact the observed associations between this association of interest, among pregnant women followed from presentation for antenatal care, through pregnancy and up to 12 months postpartum. This research will generate stronger evidence for hypothesised associations between maternal infections in general, and HIV specifically, with adverse birth outcomes using prospectively collected longitudinal data with robust measures. Evidence will also be produced that can be used to inform future research methodology when investigating birth outcomes in order avoid bias that can affect study inferences. Finally, this research will also help inform which analytical techniques are best to mitigate bias that cannot be avoided during data collection.

The specific objectives of this thesis are[3]:

1. To identify the important methodological (selection and measurement bias) issues in the epidemiological study of preterm birth and its sequelae in the context of maternal HIV and its treatment in LMICs.
2. To examine the impact of measurement error related to gestational age assessment methods on the association between maternal ART use and preterm delivery.
3. To examine the association between maternal ART use, timing of ART initiation and adverse birth outcomes, with bias adjusted gestational age
4. To describe the natural history of blood pressure throughout pregnancy and assess whether this mediates the association between ART use and adverse birth outcomes.
5. To determine whether different data sources in the same underlying population impact the associations between maternal ART use and preterm delivery.

[3] Throughout the thesis, women are described as being either living with HIV or living without HIV, with the exception of the title of Chapter 6, the tables in Chapters 4-7 and supplementary tables in Chapter 9 where they are described as being HIV-infected and HIV-uninfected.

1.4 Data sources

To address the aim and objectives of the proposed PhD research, secondary analyses were conducted using a combination of data collected prospectively through two studies: PIMS (Prematurity Immunology in HIV-infected Mothers and their Infants Study (PIMS) and B-Positive: A population based evaluation of expanded ART access in pregnancy; routine clinical record data, routine electronic medical records from the Pregnancy Exposure Register (PER) and the Provincial Health Data Centre (PHDC) of the Western Cape Department of Health, obtained retrospectively.

For all women enrolled in the PIMS cohort, the data sources provided different levels of data that was either collected as part of study procedures or as part of routine health services . Data collected during the PIMS study were used for Objectives 2, 3 and 4, and this was combined with parallel routine data from the PHDC. Objective 5 used data from the B-Positive study; as well as routine data collected during the study enrolment period from the PER for all pregnant women accessing antenatal care from the Gugulethu Midwife Obstetric Unit and PHDC for all pregnant women in the Western Cape Province. These data sources and their contributions are described in detail in Table 1.1.

1.4.1 PIMS Study

1.4.1.1 Overview

The Prematurity Immunology in HIV-infected Mothers and their Infants Study was a prospective observational cohort study which sought to investigating the association between maternal ART use and adverse birth outcomes. This study was conducted at a large public sector primary care facility Gugulethu Midwife Obstetric Unit (GMOU) in Cape Town, South Africa. Between April 2015 and October 2016, consecutive pregnant women, regardless of their HIV status, seeking antenatal care at GMOU were enrolled into the study and followed up until May 2018.

The overall prospective, observational study design included two 'nested' groups of pregnant women. Following clinical assessment, all women who agreed to enrol in the study had their routinely collected last menstrual period (LMP)-based gestational age (GA) and symphysis-fundal height (SFH)-based gestational age reviewed by the counsellor (Group 1, n=3972); women estimated to be ≤24 weeks clinically were referred for a same day ultrasound by a research sonographer using standardised assessment protocols and blinded to the midwife GA assessment. Women who identified as living with HIV and ≤24 weeks' gestation on US were recruited into a nested cohort

(Group 2, n=551); approximately half of these had initiated life-long ART prior to conception (n=261) and half ART during pregnancy (n=290).

Study participants in Group 1 were followed through pregnancy until delivery, while the subset enrolled into Group 2 were followed up longitudinally until 12 months postpartum. All research activities were conducted by research staff who were separate from routine clinical care which was managed as usual by the clinic. Group 1 participants had data collected through data abstraction from routine clinical records through the pregnancy period. Information abstracted included HIV status, previous pregnancy history, results of routine blood tests conducted during pregnancy (haemoglobin, ABO blood group, rhesus and syphilis screening), any medications prescribed during pregnancy and any maternal diagnoses during pregnancy. Obstetric outcomes, which included date of delivery and birthweight, were abstracted from clinical records at delivery. In addition to this data, at all study visits Group 2 participants had data collected on maternal health (HIV care and ART use, clinical care and inter-current clinical history), and through other procedures which included physical examinations (anthropometry, blood pressure measurement), phlebotomy; a follow-up ultrasound was conducted at 28 weeks to assess fetal growth.

Figure 1.5 presents a schema of these phases as well as the sample sizes into each Group. Different sub-samples are included in different analyses of this thesis, as described in each chapter.

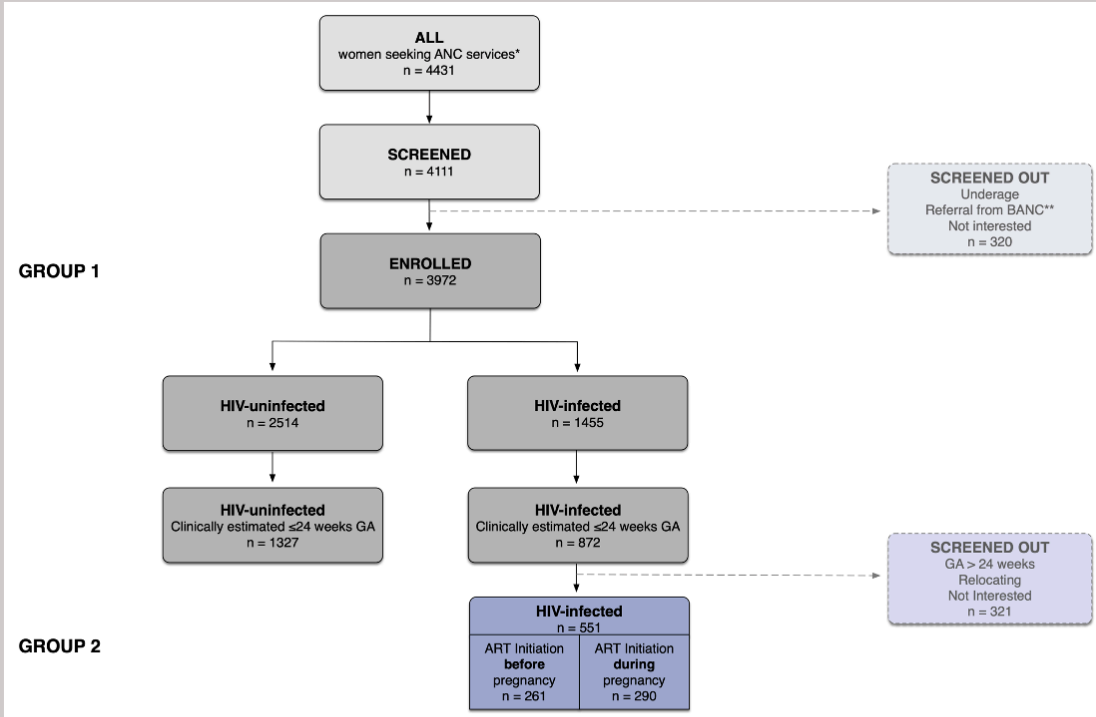

Figure 1.5 Schema of PIMS study

* 1st antenatal care visit at Gugulethu Midwife Obstetric Unit
** BANC – Basic Antenatal Care Clinic

1.4.1.2 Contribution of data to thesis

Women who enrolled into PIMS study were included in the analyses presented in Objectives 2, 3, 4 of this thesis. Gestational age, assessed by routine care staff and recorded in maternity case record (MCR) booklets during the 1st ANC visit, in conjunction with ultrasound determined gestational age by the researcher sonographer, was used to categorise gestational age based outcomes in all PIMS analyses. Explanatory variables such as obstetric history (parity, gravidity, previous pregnancy history), body mass index (BMI) and HIV-related parameters collected at the 1st ANC visit were used for PIMS analyses. Pregnancy outcomes (pregnancy loss or live birth) were either self-reported or obtained from maternity case record abstraction at delivery. Blood pressure measurements taken at the 1st ANC visit, all subsequent ANC visits and abstracted at delivery was used to examine blood pressure trajectories in pregnancy for Objective 4.

1.4.2 B-Positive Study

1.4.2.1 Overview

B-Positive was a population-level study, which sought to explore the population-level impact of the universal initiation of lifelong ART policy for pregnant women living with HIV based on strengthening routine data systems. The study was based on a combination of province-wide surveillance through harmonisation of existing routine data systems for HIV/AIDS, ART and maternal and child health, and dedicated sentinel studies in a single geographic area to investigate population-level trends in greater detail (Figure 1.6). The sentinel site component of the study explored the impact of increased ART exposure antenatally and around conception on maternal, perinatal and child health outcomes.

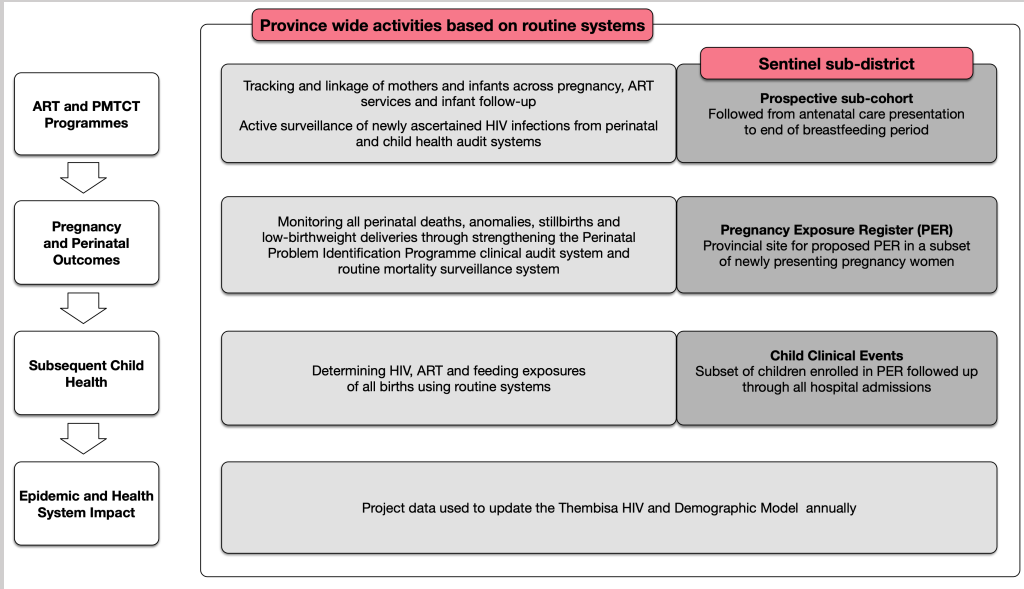

Figure 1.6 Schematic outline of study activities at both provincial and sentinel site level

Adapted from B-Positive Protocol, courtesy of Andrew Boulle

GMOU was the designated sentinel site for the B-Positive study, and this site participated in the pregnancy exposure registry (PER). For this all women attending their 1st ANC visit at the GMOU were enrolled and followed to delivery (approx. 5000/year), with their routine operational obstetric and medical data digitised from the clinical stationery, namely the patient-held MCR. Additional data sources included primary-care dating ultrasound reports, and clinic registers (STI and delivery). Nested within these total annual ANC bookings at GMOU that were included in the PER, the prospective cohort component of B-Positive (sentinel sub-cohort) was established for a more detailed cohort analysis of pregnancy and infant outcomes (Figure 1.7). Between January 2017 and July 2018, consecutive pregnant women, regardless of their HIV status, seeking antenatal care at GMOU were enrolled and followed up until March 2020.

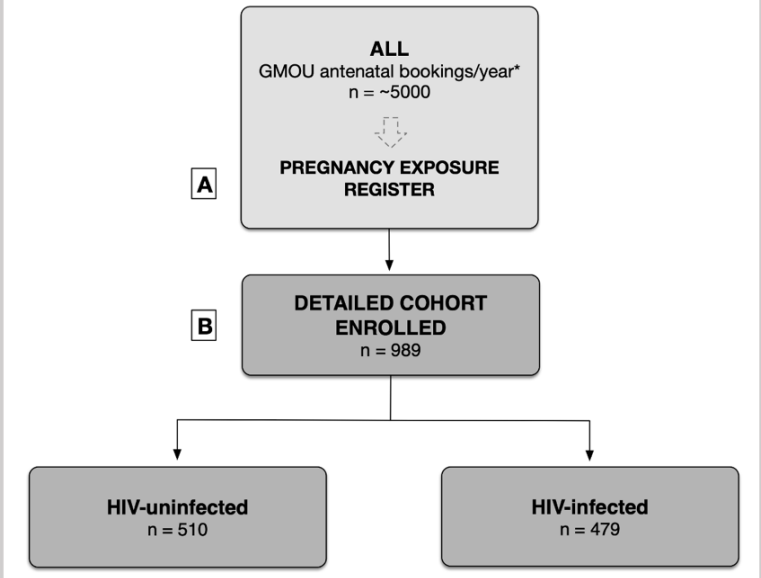

Figure 1.7 Schematic outline of B-Positive sentinel site cohorts.

A: Pregnancy Exposure Register (GMOU)
B: Detailed cohort of women booking at GMOU for antenatal care

All women enrolled in the cohort study had their gestational age determined by ultrasound with a research sonographer using standardised assessment protocols (n=989). Approximately half of the enrolled women identified as living with HIV (n=479) and the other half living without HIV (n=510). Study participants were followed up through pregnancy until 12 months postpartum. Similar to the PIMS study, all research activities were conducted separate from routine clinical care. Data was collected at up to three antenatal study visits (depending on gestational age at enrolment), and four post-natal study visits. The measurements included questionnaires on basic demographic information, maternal health (obstetric and medical history, clinical care, and inter-current clinical history), dietary intake, mental health measures and risk behaviours. Other procedures conducted included physical examinations (anthropometry, blood pressure measurement), phlebotomy and abstraction of clinical data on routine antenatal, postnatal and ART services received. Obstetric outcomes were obtained by a combination of self-report and abstraction of clinical records. Different sub-samples were included in the different components of Objective 4 analyses of this thesis, as described in each chapter.

1.4.2.2 Contribution of data to thesis

Women who were enrolled into the PER and the cohort study component of B-Positive were included in the analyses conducted for Objective 5. PER used operational data, so variable definitions were aligned with operational clinical definitions in the Western Cape Province (52). Gestational age data stored in the PER, which was used to categorise gestational age based outcomes, was based on a combination of primary-care dating ultrasound reports and clinical assessment conducted by routine care staff during the 1st ANC visit. Medication usage, including ART, which was obtained from either the clinical consultation and ward-stock medicines recorded in the MCR or electronic dispensing data from public sector pharmacies was used to construct the ART exposure variables (53). Other explanatory variables which could act as confounders for some outcomes were also recorded on PER, these include weight gain and substance use.

1.4.3 Provincial Health Data Centre

1.4.3.1 Overview

Numerous electronic record systems are maintained by the Western Cape Department of Health in order to manage clinic and hospital administration, pharmacy and laboratory data (54). All patients in fixed public sector services and facilities (primary care clinics, secondary and tertiary level hospitals) in the Western Cape Province are allocated a unique patient identifier, facilitating linkage of medical records between a variety of patient registration systems (birth, disease and death registers, health facility patient information systems (visits and admissions), National Health Laboratory Service and electronic pharmacy management systems) (53-55). This unique identifier is leveraged by the PHDC, a patient-level interlinked health information exchange housed within the WCDOH, to centrally combine various routine electronic clinical, clerical and administrative database platforms across the province to generate an individual-level population database (53, 56) (Figure 1.8). Deduplicating algorithms are used by the PHDC for identification of patients with more than one provincial identifier (53, 54, 56).

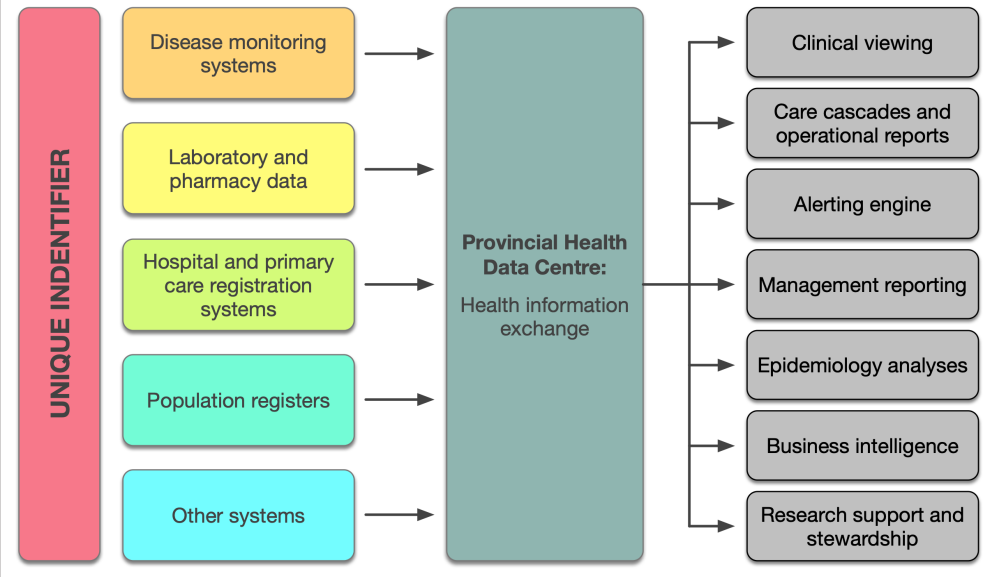

Figure 2 A schematic of the Western Cape Provincial Health Data Centre

Adapted from: Boulle A, Heekes A, Tiffin N, Smith M, et al. Data centre profile: the provincial health data centre of the Western Cape Province, South Africa. *Int J Popul Data Sci.* 2019;4(2):1-11.

1.4.3.2 Contribution of data to thesis

Laboratory, pharmacy and birth outcome data were requested from the PHDC for all women who had enrolled into PIMS. This provided routine care data from between the 1st ANC visit and delivery, which was used in conjunction with study data (questionnaires and clinical record data abstraction), especially for Group 1 women who were only followed up passively through clinical records after the 1st ANC visit. Available data for births outcomes included outcome (livebirth, stillbirth, early neonatal death), delivery method, infant date of birth, birthweight and gender)(routinely captured by into delivery registers). Pharmacy dispensing data from the local pharmacy administrative databases was used to determine ART regimens; laboratory test data were used to obtain missing CD4 cell counts and HIV viral loads from the NHLS database. These data were used to address Objectives 2, 3 and 4. For Objective 5 where birth outcomes were compared between data sources, PHDC data was used for the total population of pregnant women in the Western Cape Province over the study enrolment period. Data requested included HIV/ART status, basic obstetric history (gravidity, parity, previous adverse birth outcome), risk factors (infections, substance use, mental health), gestational age at delivery, pregnancy outcome, delivery mode and infant birth characteristics (gender, birthweight, length and head circumference).

1.5 Ethical approvals

The PIMS study protocol, informed consent documents and data collection tools have been approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (UCT-HREC; REF: 739/2014) and the University of Southampton Faculty of Medicine Ethics Committee (REF: 12542 PIMS). Study progress was reviewed annually and the latest annual approval is provided in Appendix 9.1.1. The larger B-Positive protocol (REF: 749/2015, Appendix 9.1.2) and the sentinel site study protocol (REF: 541/2015, Appendix 9.1.3) for PER and Cohort study activities together with informed consent documents and data collection tools were also approved by the UCT HREC. Additionally, permission to conduct the PIMS study (REF: WC_2015RP32_631) and the cohort component of B-Positive (REF WC_2016RP6_286) was also granted by the research oversight body of the Provincial Government of the Western Cape Department of Health.

Because of the need for parental consent for younger women, both PIMS and the B-Positive cohort studies only enrolled women aged 18 years and older. All women enrolled in the PIMS and the B-Positive cohort study provided written informed consent which included consent for the abstraction of routine clinical records, as well as linkage to electronic provincial databases. To facilitate data abstraction for these studies, identifying information was collected which included the unique provincial patient identification number, name, surname, date of birth and national identification number if available. Linked data was obtained from the PHDC for available data sources for all women who enrolled in PIMS and the B-Positive cohort study. The secondary data analyses undertaken for this thesis were approved by the UCT HREC (REF: 019/2020, Appendix 9.1.4).

1.6 Overview and structure of the thesis

This thesis consists of this introductory chapter, a literature review chapter, four results chapters, a discussion chapter and an appendix section with supporting supplemental materials. This introductory chapter contextualises the burden of child and neonatal mortality, particularly in sub-Saharan Africa; and reviews the contribution of maternal infections related adverse birth outcomes to this. Additionally, this chapter provides a brief overview of the methodological issues, related to birth outcomes, which this thesis aims to address together with the rationale, aim and objectives is included. Since this thesis involves secondary analysis, a description of the broader studies and data sources from which the data arise is included in this chapter.

Chapter 2 is a literature review that provides the background to this thesis, highlighting definitions and issues that pertain to the epidemiological measurement of pregnancy outcomes within the context of maternal ART use. This is not intended to be an exhaustive review but rather to present key aspects of birth outcomes and methodological concerns in their investigation. Details regarding the focus of this review are provided at the beginning of the chapter. Four results chapters are included (Chapters 4 to 7), which present data analyses focusing on the impact of measurement and/selection bias on the relationship between maternal ART use and adverse birth outcomes. They present results from a cohort of pregnant women accessing antenatal care at public sector health facilities in the Western Cape Province between 2014 and 2018. All results chapters are manuscripts that are either published or being prepared for submission (Table 1):

- Objective 1 is explored in Chapter 3, which highlights the important selection and measurement bias issues that need to be considered when studying the epidemiology of preterm birth and its sequelae in the context of maternal infections and their treatment in LMICs.
- Objective 2 and 3 are related to gestational age measurement error and are addressed sequentially in Chapters 4 and 5. The impact of this measurement error in the association between maternal ART use and adverse birth outcomes is established in Chapter 4. These findings provide novel data that can be used to make bias corrections of gestational age. Following this, Chapter 5 goes on to use this information to make appropriate adjustments to gestational age in the analysis in order to more accurately assess the prevalence of adverse birth outcomes and estimate the associations between HIV status, timing of ART initiation and adverse birth outcomes

- Objective 4 is covered in Chapters 6 and focuses on the measurement of maternal blood pressure using trajectories instead of cross-sectional measurements and the more detailed insight this methodology provides into the relationship between blood pressure and adverse birth outcomes in a novel analytical approach.
- Objective 5 is addressed in Chapter 7, with the results showing how using different data sources impacts the quality of exposure and outcome measure and the impact of this on the association between maternal ART use and adverse birth outcomes

Chapter 8 provides a summary of the key findings of this combined body of work, with interpretation and contextualisation of the results to identifying necessary future actions including recommendations for future research. Supplementary material including inserts from each results chapter, follows in Chapter 9.

Table 1.1 Summary of thesis results chapters, objectives, manuscripts, and data sources

Chapter and Objective		Manuscript title and status	Data Source
3	To identify the important methodological (selection and measurement bias) issues in the epidemiological study of preterm birth and its sequelae in the context of maternal HIV and its treatment in LMICs.	**Malaba TR**, Newell ML, Myer L, Ramokolo V. Methodological considerations for preterm birth research. *Front Glob Womens Health*. 2022;2:821064.	N/A
4	To examine the impact of measurement error related to gestational age assessment methods on the association between maternal ART use and preterm delivery.	**Malaba TR**, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol.* 2018; 28(12):893-900	PIMS PHDC
5	To examine the association between maternal ART use, timing of ART initiation and adverse birth outcomes, with bias adjusted gestational age.	**Malaba TR**, Mukonda E, Matjila M, Madlala HP, Myer L, Newell ML, PIMS Study Group. Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age. *Paediatr Perinat Epidemiol*. 2022;36(4):525-535.	PIMS PHDC
6	To describe the natural history of blood pressure throughout pregnancy and assess whether this mediates the association between ART use and adverse pregnancy outcomes.	**Malaba TR,** Cois A, Madlala HP, Matjila M, Myer L, Newell ML. Blood pressure trajectories during pregnancy and associations with adverse birth outcomes among HIV-infected and HIV-uninfected women in South Africa: a group-based trajectory modelling approach. *BMC Pregnancy Childbirth*. 2020;20(1):742-754.	PIMS PHDC
7	To determine whether different data sources in the same underlying population impact the associations between maternal ART use and preterm delivery.	Being prepared for submission	B-Positive PER PHDC
9	**Appendix supporting PIMS analyses**	**Malaba TR**, Myer L, Gray CM, Newell ML. Cohort profile: Prematurity Immunology in Mothers living with HIV and their infants Study (PIMS). *BMJ Open*. 2021;11(9): e047133.	N/A

1.7 References

1. UNAIDS. Fact Sheet: World AIDS Day 2021. Geneva, Switzerland: United Nations Joint Program on HIV/AIDS; 2022.
2. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: summary of key features and recommendations. Geneva, Switzerland: World Health Organization; 2013.
3. Gilleece Y, Krankowska D. ART in pregnant women living with HIV. *Lancet*. 2021;397(10281):1240-1241.
4. Perin J, Mulick A, Yeung D, et al. Global, regional, and national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet Child Adolesc Health*. 2022;6(2):106-115.
5. Sharrow D, Hug L, You D, et al. Global, regional, and national trends in under-5 mortality between 1990 and 2019 with scenario-based projections until 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *Lancet Glob Health*. 2022;10(2):e195-e206.
6. UN-IGME. Levels and Trends in Child Mortality 2021. New York, USA: United Nations Inter-agency Group for Child Mortality Estimation; 2021.
7. UN-IGME. Levels and Trends in Child Mortality Report 2018. New York, USA: United Nations Inter-agency Group for Child Mortality Estimation; 2018.
8. UN. The Millenium Development Goals Report. New York, USA: United Nations; 2015.
9. WHO. Millennium Development Goals: Goal 4: Reduce Child Mortality. Geneva, Swizterland: World Health Organization; 2006.
10. You D, Hug L, Ejdemyr S, et al. Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the United Nations Inter-agency Group for Child Mortality Estimation. *Lancet*. 2015;386(10010):2275-2286.
11. WHO. Global Action Plan: For health lives and well-being for all. Geneva, Switzerland: World Health Organization; 2018.
12. Yovo E. Challenges on the road to achieving the SDG 3.2 targets in resource-limited settings. *Lancet Glob Health*. 2022;10(2):e157-e158.
13. UN. Sustainable development goal 3: ensure healthy lives and promote well-being for all ages. New York, USA: United Nations; 2020.

14. Lassi ZS, Middleton PF, Crowther C, Bhutta ZA. Interventions to improve neonatal health and later survival: an overview of systematic reviews. *EBioMedicine*. 2015;2(8):985-1000.
15. Li Z, Karlsson O, Kim R, Subramanian SV. Distribution of under-5 deaths in the neonatal, postneonatal, and childhood periods: a multicountry analysis in 64 low- and middle-income countries. *Int J Equity Health.* 2021;20(1):109-120
16. Hug L, Alexander M, You D, Alkema L, for UN Inter-agency Group for Child Mortality Estimation. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *Lancet Glob Health*. 2019;7(6):e710-e720.
17. Katz J, Lee AC, Kozuki N, et al. Mortality risk in preterm and small-for-gestational age infants in low-income and middle-income countries: a pooled country analysis. *Lancet.* 2013;382(9890):417–425.
18. Christian P, Lee SE, Donahue Angel M, et al. Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low-and middle-income countries. Int J Epidemiol. 2013;42(5):1340-1355.
19. Liu L, Johnson HL, Cousens S, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet.* 2012; 379(9832):2151-2161
20. UN-IGME. Levels and Trends in Child Mortality: Report 2019. New York, USA: United Nations Inter-agency group for child mortality estimation; 2019.
21. Lee AC, Kozuki N, Cousens S, et al. Estimates of burden and consequences of infants born small for gestational age in low- and middle-income countries with INTERGROWTH-21st standard: analysis of CHERG datasets*. BMJ.* 2017;358:3677-3688.
22. Ludvigsson JF, Lu D, Hammarström L, Cnattingius S, Fang F. Small for gestational age and risk of childhood mortality: A Swedish population study. *PLoS Med.* 2018;15(12):e1002717.
23. Bukowski R, Hansen NI, Willinger M, et al. Fetal growth and risk of stillbirth: a population-based case–control study. *PLoS Med*. 2014;11(4):e1001633.
24. Paixao ES, Blencowe H, Falcao IR, et al. Risk of mortality for small newborns in Brazil, 2011-2018: A national birth cohort study of 17.6 million records from routine register-based linked data. *Lancet Reg Health Am*. 2021;3:100045.
25. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet*. 2012;379(9832):2162-2172.
26. Lee AC, Katz J, Blencowe H, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Glob Health*. 2013;1(1):e26-36.

27. Blencowe H, Krasevec J, de Onis M, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *Lancet Glob Health*. 2019;7(7):e849-e860.

28. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet. 2008; 371(9608):261-269

29. Kim H, Cizmadia P. Adverse Birth Outcomes and Environmental Health Threats. Physicians for Social Responsibility. https://www.psr.org/assets/pdfs/abo-fact-sheet.pdf (accessed 20 December 2018)

30. Weng YH, Yang CY, Chiu YW. Risk assessment of adverse birth outcomes in relation to maternal age. *PLoS One*. 2014;9(12):e114843.

31. Kramer MS, Victora CG. Low birth weight and perinatal mortality. In: Semba RD, Bloem MW (Eds). Nutrition and health in developing countries. pg 57–69. Humana Press: New Jersey, USA; 2002.

32. Clausson B, Cnattingius S, Axelsson O. Preterm and term births of small for gestational age infants: a population-based study of risk factors among nulliparous women. *Br J Obstet Gynaecol*. 1998;105(9):1011-1017.

33. Sania A, Spiegelman D, Rich-Edwards J, et al. The Contribution of Preterm Birth and Intrauterine Growth Restriction to Infant Mortality in Tanzania. *Paediatr Perinat Epidemiol*. 2014;28(1):23-31.

34. Brocklehurst P, French R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *Br J Obstet Gynaecol*. 1998;105(8):836-848.

35. Kapisi J, Kakuru A, Jagannathan P, et al. Relationships between infection with Plasmodium falciparum during pregnancy, measures of placental malaria, and adverse birth outcomes. *Malar J*. 2017;16(1):400-411.

36. Qin J, Yang T, Xiao S, et al. Reported estimates of adverse pregnancy outcomes among women with and without syphilis: a systematic review and meta-analysis. *PLoS One*. 2014;9(7):e102203.

37. Watson-Jones D, Weiss HA, Changalucha JM, et al. Adverse birth outcomes in United Republic of Tanzania: impact and prevention of maternal risk factors. *Bull World Health Organ*. 2007;85(1):9-18.

38. Weckman AM, Ngai M, Wright J, McDonald CR, Kain KC. The impact of infection in pregnancy on placental vascular development and adverse birth outcomes. *Front Microbiol*. 2019;10(1924):1-11.

39. Conroy AL, McDonald CR, Gamble JL, et al. Altered angiogenesis as a common mechanism underlying preterm birth, small for gestational age, and stillbirth in women living with HIV. *Am J Obstet Gynecol*. 2017;217(6):684. e1- e17.
40. Chua CL, Hasang W, Rogerson SJ, Teo A. Poor birth outcomes in malaria in pregnancy: recent insights into mechanisms and prevention approaches. *Front Immunol*. 2021;12(621382):1-11.
41. Sania A, Smith ER, Manji K, et al. Neonatal and infant mortality risk associated with preterm and small for gestational age births in Tanzania: individual level pooled analysis using the intergrowth standard. *J Pediatr.* 2018;192:66-72.
42. Zash R, Jacobson DL, Diseko M, Mayondi G, Mmalane M, Essex M, et al. Comparative safety of antiretroviral treatment regimens in pregnancy. *JAMA Pediatr*. 2017;171(10):e172222-e.
43. Snowden JM, Bovbjerg ML, Dissanayake M, Basso O. The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection. *Curr Epidemiol Rep*. 2018;5(4):379–387.
44. Neophytou AM, Kioumourtzoglou M-A, Goin DE, Darwin KC, Casey JA. Addressing special cases of bias that frequently occur in perinatal epidemiology*. Int J Epidemiol*. 2021;50(1):337-345.
45. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, USA; 2008.
46. Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic measures of the course and outcome of pregnancy. *Epidemiol Rev.* 2002;24(2):91-101.
47. Hernán MA, Schisterman EF, Hernández-Díaz S. Invited commentary: composite outcomes as an attempt to escape from selection bias and related paradoxes. *Am J Epidemiol*. 2013;179(3):368-370.
48. Idele P, Hayashi C, Porth T, Mamahit A, Mahy M. Prevention of mother-to-child transmission of HIV and paediatric HIV care and treatment monitoring: from measuring process to impact and elimination of mother-to-child transmission of HIV. *AIDS Behav.* 2017;21(1):23-33.
49. Arbogast PG, VanderWeele TJ. Considerations for statistical analysis. In: Velentgas P, Dreyer NA, Nourjah P, Smith SR, Marion M Torchia MM (Eds). Developing a protocol for observational comparative effectiveness research: a user's guide. Agency for Healthcare Research and Quality: Maryland, USA; 2013.
50. Groenwold RH, Palmer TM, Tilling K. To adjust or not to adjust? When a “confounder” is only measured after exposure. *Epidemiology*. 2021;32(2):194-201.
51. Nannine N, Dorrington R, Laubscher R. Under-5 mortality statistics in South Africa. South African Medical Research Council. Cape Town, South Africa; 2010.
52. Kalk E, Metha U, Slogrove AL, et al. Pregnancy Exposure Registry/Birth Defects Surveillance Programme In The Western Cape, South Africa: A Model For Low- & Middle-Income Countries. *18th Annual meeting International Society of Pharmacovigilance*, Geneva, Switzerland, 2018.

53. Mehta U, Heekes A, Kalk E, Boulle A. Assessing the value of Western Cape Provincial Government health administrative data and electronic pharmacy records in ascertaining medicine use during pregnancy. *S Afr Med J*. 2018;108(5):439-443.

54. Heekes A, Tiffin N, Dane P, et al. Self-enrolment antenatal health promotion data as an adjunct to maternal clinical information systems in the Western Cape Province of South Africa. *BMJ Glob Health*. 2018;3(Suppl 2):e000565.

55. Beck EJ, Shields JM, Tanna G, et al. Developing and implementing national health identifiers in resource limited countries: why, what, who, when and how? *Glob Health Action*. 2018;11(1):1440782.

56. Boulle A, Heekes A, Tiffin N, et al. Data centre profile: the provincial health data centre of the Western Cape Province, South Africa. *Int J Popul Data Sci*. 2019;4(2):1-11.

Chapter 2: Literature Review

2.1 Overview

The purpose of this chapter is to establish the conceptual background for this thesis by providing an orientation to and summary of existing knowledge pertaining to the epidemiological measurement of outcomes that occur over the course of pregnancy (conception to birth and the immediate neonatal period). Important considerations related to pregnancy and perinatal outcome assessments will be highlighted. Particularly as these outcomes are primarily defined according to their timing of occurrence or detection, which can either be *in utero* or in the perinatal period. The occurrence of adverse outcomes is influenced by *in utero* events and exposures, so this chapter also summarises key perinatal epidemiology concepts related to measurement and selection issues that may impact findings of associations between exposures and adverse outcomes. The latter part of this chapter is structured according to the objectives of this thesis and address's themes that are relevant to the results chapters which follow. Specifically, given the importance of maternal infections and their treatment in the occurrence of adverse birth outcomes in Sub-Saharan Africa, these concepts will be explored within the context of maternal antiretroviral therapy (ART) use in pregnancy, which is the main exposure of interest in this thesis.

This review is not intended to be systematic or exhaustive but rather provides an overview of key points and gaps in our understanding of the methodological considerations when investigating outcomes of pregnancy. While discussions of the impact of the highlighted methodological considerations on adverse birth outcomes are related to maternal ART use in pregnancy, this review also includes a few articles from epidemiological studies of other maternal exposures, infections and/or treatments where these contribute important knowledge. Similarly, the literature reviewed is primarily from low to middle income countries (LMIC), since the data on which this thesis is based arise from South Africa, however literature from high-income countries (HIC) is included if it provides a meaningful contribution. PubMed was searched for the relevant literature, and reference lists were checked for additional key publications.

Section 2.2 presents an overview of the normal events that occur during the reproductive process through a discussion of the measures of pregnancy course and outcomes. Problems that arise during this reproductive process define the adverse outcomes in epidemiologic studies of pregnancy, so this section will address the deviations from normal events that occur that result in adverse outcomes. The focus will be on the main outcomes of interest in this thesis, namely those related to gestational duration (preterm and post-term delivery) and fetal growth (birthweight and

size for gestational age). Although not the primary focus of this thesis, this section also covers pregnancy losses that can occur during the course of pregnancy, ranging from early pregnancy loss to miscarriage and stillbirths. While congenital anomalies are an important adverse outcome of pregnancy, they are not covered in this section, as a separate outcome, because they present a different set of methodological considerations which fall outside the scope of this thesis.

Section 2.3 provides a broad overview of important methodological issues which require careful consideration when investigating birth outcomes. The purpose of this section is not to provide an extensive review of all methodological issues that apply in perinatal epidemiology but rather to briefly describe issues related specifically to outcomes related to gestational duration, fetal growth and pregnancy loss. In addition to these broader concerns, each of the outcomes have their own unique challenges linked to study design choices, outcome assessment and potential threats to validity that are also reviewed in this section.

Section 2.4 reviews literature to provide context for the main objectives of this thesis, related to the specific cases of bias that are often encountered in perinatal epidemiology. An overview of information bias is covered with a focus on outcome measurement error and misclassification, and potential solutions of different types of correction methods, providing a strong basis for the methodological approaches taken in this thesis. Selection bias is also covered with a focus on the specific types that are relevant to the study of birth outcomes, with a brief discussion of mitigation strategies.

Section 2.5 contextualises the main area of interest in this thesis, the relationship between maternal ART use and adverse birth outcomes. The data on which this thesis is based on are recent data from South Africa, so more weight is given throughout this literature review, to relatively recent studies in similar settings with a focus on triple drug antiretroviral therapy. Finally, this section highlights bias related concerns of the current evidence, and points to the future research needs that the subsequent results chapters aim to address.

2.2 Definitions and measures of pregnancy course and outcomes

Successful reproduction is biologically and epidemiologically complex; because pregnancy comprises discrete events which range from implantation, decidualisation, placentation and finally infant birth (1, 2). The completion of each of these events is essential for advancement toward the next stage (2). The path towards a successful pregnancy begins during implantation, approximately 7 days after conception, through the initiation of complex but highly organised and precisely timed physiological

and molecular processes (3). This is incidentally also when pregnancy can first be detected through monitoring of the human chorionic gonadotropin (hCG) hormone levels in serum and urine (4, 5).

Normal development following implantation starts with the formation of organ systems, followed by fetal growth and development (6). Organ growth affects the developmental environment, and during periods of critical organ growth irrevocable developmental decisions are made (6). There are two critical periods of fetal growth, arising from patterns of growth in cell number and size. The first critical period is the embryonic phase, which lasts from the 5th to the 10th gestational age week (7), during which organs are formed in a manner that requires each development stage occurring at the correct time (1). At approximately 4 weeks gestation the first ultrasonographically visible structure of pregnancy is the gestational sac, a fluid-filled structure that surrounds the embryo during the first few weeks of embryonic development (8). The diameter and volume of the gestational sac can be measured to determine gestational age (9). The next visible structure, at approximately 5½ weeks' gestation, is the yolk sac a circular structure within the gestational sac that provides early transfer of nutrients from the trophoblast to the embryo (before the placenta takes over this function) (10). By 6 weeks gestation, the embryo can first be identified as a focal thickening on the outer margin of the yolk sac. As the first trimester progresses the outer layer of the embryo merges with the endometrium, forming the placenta, which takes over the nutrient and waste requirements of the embryo. The size of the embryo is determined using the crown-rump length (CRL), which measures length from the top of the head (crown) to the bottom of the buttocks (rump). This measure translates to an accurate estimated gestational age in early pregnancy (11). At the end of the embryonic phase most organ systems are in place albeit in undeveloped states (7).

The second critical period of fetal growth is the fetal phase, which occurs after the completion of organogenesis at approximately 10 weeks) (7). At this stage all major organs have been formed, and the embryo has become a fetus, with distinct human features. During the first half of the fetal phase (gestational weeks 11 to 26), organ size increases rapidly, primarily through cell proliferation, and become more developed. The rate of fetal growth varies over gestation – from week 9 to 20 the CRL increases ~4-fold, while fetal weight increases ~60-fold. Subsequently the rate slows, with a 2-fold CRL increase and 5-fold fetal weight increase by delivery (10). Beyond 12 to 14 weeks' gestation the excessive curvature of the fetus may lead to erroneous shortening of CRL measurements, consequently in the second and third trimester other fetal measurements are required for gestational age assessment (12). These include the biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC) and femur length (FL). While gestational age can be assessed using a single parameter, the use of multiple parameters has been shown to be more accurate (13, 14). In the second half of the fetal period (27 weeks to term), organ cell numbers stabilise and there

is a shift to the hypertrophic stage of growth where the size of cells increase (1). After a period of relatively constant weight gain in the second trimester, absolute weight increases in the third trimester at an average fetal weight increase of 2000 grams (10), linked to the rapid accumulation of fat, muscle, and connective tissue (15). Cellular hypertrophy continues postnatally into childhood according to varying schedules for different organs. During this period of development there is also a high placental transfer of essential nutrients required to support the infant during the first 6 months of life (16).

Complications or deviations from the normal reproductive process define adverse outcomes in epidemiological studies of pregnancy. Adverse outcomes of pregnancy are either due to the failure of a developmental process to occur at the correct gestational time point, or the occurrence of normal developmental processes at the incorrect gestational time point. Additionally, adverse outcomes may also follow disruptions during critical periods of development by potentially harmful exposures such as infections, medications, and other teratogenic substances. During the first trimester when major structures are formed during organogenesis these disruptions or harmful exposures have the greatest chance of causing congenital anomalies. During the second and third trimester where growth is the focus, harmful exposures can result in growth retardation and minor congenital anomalies and can be associated with preterm delivery (PTD). The use of epidemiology to identify promising approaches to prevention or development of appropriate interventions requires elucidation of the causes of these adverse pregnancy outcomes. However, to discern causal determinants and patterns of associations with predictors and adverse pregnancy outcomes, consideration of the underlying methodological nuances of reproductive and perinatal epidemiology is of paramount importance. These are linked to choice of study population and appropriate units of analysis, interval censoring and truncation, and hidden or missing data (6, 17), which will be discussed in Section 2.3 of this review.

Although the study of pregnancy includes the determinants of conception and postpartum health and development through to infancy, childhood and adulthood, this literature review will be restricted to the timeframe between conception and birth and deviations from normal events that lead to adverse outcomes specifically related to gestational duration and fetal growth (Figure 2.1).

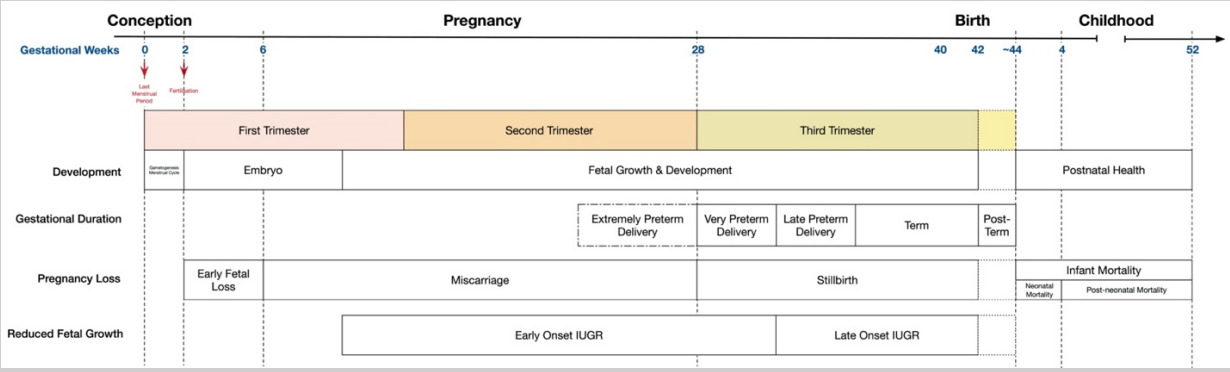

Figure 2.1 Overview of the timeline of pregnancy-related outcomes covered

Adapted from: Slama R. Epidemiologic Tools to Study the Influence of Environmental Factors on Fecundity and Pregnancy-related Outcomes. *Epidemiol Rev.* 2014;36(1):148-164.

2.2.1 Gestational Duration

Fetal maturation, the process of physiological development, prepares for fetal survival outside the uterus (18). Gestational duration is the most accessible measure of fetal maturation and is important for several epidemiological endpoints. This is because at any given week of gestation, any pregnancy can be interrupted. An example of one such interruption is placental abruption, which in turn can trigger labour, and, depending on the severity and gestation when this occurs, lead to preterm delivery, stillbirth and/or perinatal death. Since the initiating event is the same, the difference in outcomes can be attributed to the gestational duration at the time of the event (19); and therefore, gestational duration is arguably the most important measure in reproductive and perinatal epidemiology because it enables outcome classification. Birthweight is the final product of intrauterine growth rate over the duration of gestation, affected by any interruption to either growth rate or pregnancy duration. However, of these two outcomes, gestational age is a better short- and long-term outcome predictor than birthweight, which does not consider the appropriateness for gestational age.

Gestational age is the strongest single predictor of perinatal outcomes because of its strong correlation with all developmentally regulated processes (20) – with fetal maturity directly related to the duration of gestation. Outcome measures of gestational age are defined by duration of competed gestation at the time of delivery expressed in completed weeks or days: preterm, term and post-term. Preterm delivery accounts for the majority of perinatal morbidity, mortality, and costs (21), but post-term delivery is also associated with increased neonatal morbidity and mortality (22, 23). Clinically, information on gestational age is critical because accurate interpretation of observations, such as fetal size and assessment of appropriateness of intrauterine growth (through birthweight as a proxy), require knowledge of the duration of completed gestation at the time of measurement. In addition, clinical decisions in terms of management relate to gestational age, with

options depending on whether the fetus is delivered too early (preterm), at term, or too late (post-term).

Despite its central importance, gestational age in individual pregnancies can only be estimated approximately, and this imprecision has significant consequences for clinical care and pregnancy research. Ideally, gestational age would be measured from the day of conception, however since natural conception is unobservable, this can only be determined through hormonal markers of ovulation. These include detecting the luteinising hormone surge preceding ovulation or the sudden changes in oestrogen and progesterone following ovulation (24, 25). While this data is observable it is often not available, consequently the less direct standard clinical benchmark for the onset of a pregnancy - the first day of the last menstrual period (LMP) is used, which is on average 14 days before conception. Using this estimate, gestational duration corresponds to the time from the first day of the LMP to delivery, with an estimated median duration of 40 gestational weeks or 280 days (26). The reliability of this measure depends on several factors which other than accurate LMP recall, include cycle regularity, presence of early or light bleeding and factors that could influence ovulation timing; consequently, even with the most accurate LMP recall there is still a margin of error which could be quite substantial (27).

Gestational age can also be estimated early in pregnancy using fetal size – since it is assumed at this stage in pregnancy, fetal size is a function of age and not growth rate. The earlier the ultrasound (US) measurements are obtained the more accurate gestational age estimates will be, because of the limited variability in fetal size in early gestation. Ultrasound-based gestational age estimates may be used in conjunction with LMP estimates, if deemed appropriate (28). However, in resource-limited settings US facilities are generally unavailable in public facilities. This together with low literacy rates, high hormonal contraception rates, and nutritional challenges which contribute to unreliable LMP dates, mean gestational age assessment typically relies on other antenatal and postnatal assessment methods. Antenatally measurement of the symphysis fundal height (SFH) is a commonly practiced non-invasive method of assessing fetal growth using a tape measure (29). Like US, SFH variations in fetal size at a given gestational age can be converted into differences in gestational age. For fetuses growing normally from 24 weeks gestational age, SFH measurements in centimetres correspond to gestational weeks. Postnatal assessment of the newborn gestational maturity is based on external (developmental morphology) and/or neurological criteria (30-33); these methods are particularly useful in cases where antenatal care (ANC) was either accessed late in pregnancy or not at all.

Outside of measurement error related to assessment method used, considerable biological variation exists in the duration of healthy term pregnancies, with timing of delivery depending on the natural course of pregnancy and clinical management (1). This reflects normal variation in the pace of fetal maturation and timing of natural delivery with differences in maternal capacity to carry the fetus as term approaches. Multiple pregnancies are likely to end earlier than singleton pregnancies due to the effects of increased physical load in the uterus (6). Over time, distributions of gestational age at delivery have shifted towards early gestational ages with the recent downward shift in gestational age distribution, especially in high and middle-income countries, likely a reflection of an increase in obstetric interventions. This is due to clinicians intervening to induce labour and delivery before it occurs naturally, related to concerns around maternal and fetal wellbeing and less well-defined social factors such as clinician and maternal convenience (34).

2.2.1.1 Preterm Delivery

Preterm delivery is defined as delivery before 37 completed weeks of gestation (6). More stringent cut-offs have been applied to define more extreme deviations from the normal 37-41 week pregnancy. This is motivated, in part, by advancements in the quality and efficacy of neonatal care with significantly lowered morbidity and mortality risk in marginally preterm infants. Preterm delivery is recognised as a major public health problem because it is the leading cause of infant mortality, with the risks for survival and morbidity increasing with delivery at gestational ages below 37 weeks (35). This risk increases more steeply as the gestational duration becomes progressively shorter (36), consequently cut-offs for very preterm (<32 weeks) and extremely preterm (<28 weeks) define severe adverse outcomes. Preterm delivery reflects a final mutual pathway for numerous different clinical and biological processes resulting from a variety of causal mechanisms and risk factors (17). The PTD entity is abnormal in that it is defined by time and not a distinguishing clinical phenotype (37), so it requires closer consideration and possible subgrouping in order to study both aetiology and prognosis. Consequently, in addition to classification by gestational age, PTD can also be phenotypically classified according to clinical presentation distinguishing spontaneous PTD from medically indicated PTD (37). These two PTD phenotypes have considerable aetiological heterogeneity that require distinct prevention and management strategies (35, 38). Spontaneous PTD can be due to preterm labour with cervical dilation or preterm prelabour rupture of membranes. In contrast medically indicated PTD are those in which concerns around the health of the pregnancy lead to a clinical judgement that labour should be induced, frequently due to maternal and fetal factors such as preeclampsia and fetal distress (35, 39). Within the spontaneous PTD phenotype, further classifications have been developed to enable identification of more specific aetiological pathways. These pathways include infection/inflammation, decidual haemorrhage, maternal stress, cervical insufficiency, uterine distention, placental dysfunction, preterm premature

rupture of membranes, maternal co-morbidities, and familial factors (40). This classification is better suited to well-resourced settings or research studies because of the human and infrastructural capacity required for its application.

2.2.1.2 Post-term Delivery

Post-term delivery is defined as delivery that occurs after 41 completed weeks of gestation (41). Post-term delivery is not associated with a well-defined physiological condition (42). Advancing gestation increases the risk of conditions such as placental insufficiency, fetal acidemia and preeclampsia (17). Post-term fetuses tend to be larger than term fetuses, increasing the risk of intrapartum complications, and also associated with considerable mortality and long-term morbidity (42), with gestational age-specific mortality increasing at week 42 onwards after a period of low mortality risk between 39-41 weeks (1). However, the direction and cause of post-term related mortality is unclear – either these prolonged pregnancies increase risk of fetal complications or fetuses with complications fail to trigger delivery at the appropriate time (1).

Post-term delivery is most commonly assumed to be because of errors in gestational age assessment. Reliance on standard clinical criteria to estimate the estimated date of delivery tends to overestimate gestational age (23, 42). While inaccurate recall of LMP is known to contribute to inaccurate gestational age, irregular menstrual cycles which deviate from the norm of 28-day cycles also contribute to these inaccuracies (17). In any given cycle, ovulation is more likely to be delayed than early (43), consequently even with accurate recall of LMP gestation duration can be over-estimated with post-term pregnancy erroneously diagnosed. Other clinical methods which contribute to inaccurate gestational age include using measurement of SFH, perception of fetal movements and fetal heart tone auscultation (44).

In cases where a post-term pregnancy truly exists the cause is usually unknown because the pathogenesis is not clearly understood. While there has been improved understanding of parturition, mechanisms that initiate labour and allow its progression are also still not clear. Given that parturition involves the interaction of hormonal, inflammatory and mechanical processes in the placenta, mother and fetus, it is possible that changes in the biological mechanisms regulating delivery could lead to post-term delivery (44). These changes could be linked to hormonal factors and/or genetic predisposition (45). For example, altered levels of circulating hormones, such as placental sulphatase, that play a role in spontaneous labour could lead to post-term pregnancy (46). Genetic factors may be involved with prolongation of pregnancy, with women who were products of prolonged pregnancy at increased risk of post-term pregnancy (47).

2.2.2 Fetal Growth

Fetal growth is heavily modulated by placental function; appropriate fetal growth and development reflects fetal health, which in turn predicts infant health, development and lifelong well-being. Growth is commonly divided into three stages. The first stage occurs throughout the first trimester up to 20 weeks' gestation, with limited fetal weight gain with an average growth rate of ~5 grams/day. The second stage from 20 to 28 weeks' gestation, has an average growth rate of ~15-20 grams/day. The third stage from 28 weeks until term has the largest weight gain of ~30-35 grams/day (48). Given that fetal growth is fundamentally linked to development, any disruptions during these periods or aberration in fetal condition can result in growth failure in the form of fetal growth restriction or overgrowth. Patterns of abnormal growth are either symmetric, asymmetric or a combination of both (mixed), based on the relative sizes of organs (49).

Fetal growth restriction occurs when the growth potential of a fetus is not achieved (50), usually a result of placental insufficiency (51, 52). The consequence of this impaired placental function is its inability to provide the fetal demands for appropriate growth and development (53). While this growth restriction can occur at any time in pregnancy, the natural history of early and late onset has different biochemical, histological and clinical features. Early onset growth restriction often manifests as more severe disease and is associated with maternal hypertensive disorders and abnormal umbilical artery Doppler findings. In contrast, late onset (>32 weeks), which is harder to diagnose, is associated with a milder degree of placental dysfunction and more favourable perinatal outcomes (54). Symmetric growth restriction typically occurs relatively early in pregnancy during the hyperplastic growth phase with early disruption impacting fetal organs and their development proportionally. This suggests that growth could have been disrupted by a defect or event occurring during organogenesis, with common disruptors being congenital abnormalities, genetic disorders, and infections (7). Fetuses undergoing symmetric growth restriction have organs that are relatively small but proportional for gestational age (55, 56). Asymmetric growth restriction, which accounts for the majority of growth abnormalities, occurs later in pregnancy leading to disproportionate organ growth and development (49). This disruption of growth typically occurs during the hypertrophic growth phase when fetal growth and development are heavily dependent on sufficient nutrient and oxygen supply through the placenta (57, 58). Fetal growth restriction is a major contributor of stillbirth, neonatal morbidity and mortality and carries an increased risk of long term neurological and neurodevelopmental complications (59-61). There are indications that consequences of growth restriction later in life are associated with asymmetric rather than symmetric fetal growth restriction (62).

On the opposite end of the spectrum, a fetus can experience growth acceleration leading to excessive *in utero* growth. Fetal overgrowth or macrosomia is frequently due to maternal obesity and/or diabetes or other maternal and placental factors which affect nutrient supply and uptake to the fetus could also contribute to overgrowth (48). Patterns of fetal overgrowth can also be subdivided into symmetric and asymmetric (62). Symmetric overgrowth typically occurs when the excessive fetal weight is the result of proportionate growth of all fetal organs (63). This is usually the result of a post-term pregnancy or genetic factors such as parental anthropometric characteristics. Fetuses undergoing symmetric overgrowth have fetal parameter ratios within the normal range for gestational age and are classified as large-for-gestational-age (63). Asymmetric overgrowth is characterised by disproportionate organ growth (64), and typically originates from an abnormal intrauterine environment. This can be a consequence of maternal hyperglycaemia, which stimulates fetal insulin production (63, 65, 66). Fetal overgrowth is also associated with stillbirth and neonatal morbidity and mortality, with asymmetric associated with increased risk of neonatal complications as well as long-term health consequences (62, 67-69).

Different measures are used to identify the determinants of fetal growth, and distinctions are required between measures of fetal size and growth as these are not synonymous (70). Size is a single measure taken at a specific timepoint, with repeated measures representing distant variations in size. Conversely, growth is a dynamic process, with a change in the measure per unit of time (71, 72). Fetal growth can be measured either directly based on ultrasonographic measurement or indirectly based on anthropometric measurement at delivery. Ultrasound can establish gestational age in the first trimester, thereafter, estimated fetal weight (EFW) can be determined in relation to gestational age as an indicator of normal fetal growth and well-being (73). While AC can be used as a single parameter, combination of multiple parameters (BPD, HC, AC and FL) to determine EFW reduces potential error (74). Because growth disturbances tend to represent progressive disease, measurement at a single time point may not be able to detect any abnormal biometry that may develop. Whereas assessment over time based on the trajectory of individual parameters, provides a significant advantage as it enables evaluation of true growth parameters (growth rates) and trajectories, particularly in the third trimester when most growth abnormalities occur (75).

Anthropometric measures at delivery are also used as proxy measures of fetal growth, metrics for this include birthweight and size for gestational age. Birthweight is a reflection of the outcome of fetal growth processes and complications (17), as such it is a powerful predictor of infant's survival chances (76). An infant's birthweight is the first recorded weight after birth and should ideally be measured within the first hours of life before any significant postnatal weight loss occurs (77).

At a population level the frequency distribution of birthweight is mound-shaped with heavier tails than in a normal distribution. Deviation exists in the lower and upper residual tails, where there is an excess of low and high birthweight infants (78). The abnormal sized infants in the residual tails are at high risk of morbidity and mortality (76). Several cut-offs, ratios and indices have been devised to relate birthweight to morbidity and mortality risk (17). Birthweight alone as an endpoint measure of fetal growth does not provide sufficient information on fetal growth or growth velocity in specific gestational periods (79). Size for gestational age is a more accurate measure as it assesses birthweight in relation to measures of central tendency for gestational age. It is more useful in evaluating fetal growth outcomes and more sensitive in indicating growth abnormalities than birthweight alone. Additionally, it can help determine whether an infant is small/large because of abnormal growth and consequently at risk of morbidity or mortality; or whether the infant is constitutionally small/large and thus likely to be healthy.

The recognition of abnormal growth is dependent on the existence of reliable birthweight- and estimated fetal weight for gestational age growth charts. The statistical thresholds of these growth charts for defining 'at risk' fetuses are standardised by convention – with those falling below the 10th percentile (or to a lesser extent, the 5th or 3rd) classified as small for gestational age (SGA) and those above the 90th percentile large for gestational age (LGA) (80). There is however substantial variation in characteristics of the study populations from which these percentiles are derived. While there is recognition that appropriate unified charts are required, two different approaches have been advocated: references and standards (69). Growth references are descriptive charts of how a fetus/infant is actually growing. These references can be population-specific and are based on the premise that fetal growth can be strongly influenced by genetic factors (69, 80). They are derived from representative populations of fetuses/infants from normal and complicated pregnancies including those that may have exposures or conditions that could affect growth (17, 80, 81). In contrast, growth standards are proscriptive charts of how a fetus/infant should be growing. These standards are typically not population-specific and are based on the theory that across populations, growth potential is similar with any deviation's indicative of external influences rather than inherent biological differences (69). They are derived from selected populations of healthy low risk pregnancies, with the exclusion of cases that have exposures that may affect growth (80). Recently two methodologically rigorous growth standards have been released which were derived from multi-country populations, these include the INTERGROWTH-21st and WHO charts (75, 82, 83). Careful consideration is required when deciding between using a reference or a standard for fetal growth evaluation, given the differences in selection of source population and assumptions. Theoretically, standard charts that reflect optimal fetal growth are likely to be better for monitoring

growth and identifying 'at risk' fetuses who are experiencing abnormal growth than reference standards.

2.2.2.1 Low Birthweight

Low birthweight (LBW), defined as birthweight below 2500 grams (84), is the most common measure used in epidemiological studies and worldwide surveillance (85). Before 1976, the terms LBW and premature were used interchangeably as it was presumed infants were LBW because they were born too early (76). It was for this reason that at the time the standard criteria used to define prematurity was based on birthweight below 2500 grams regardless of period of gestation (86). Historically the LBW cut-off was used because it represented the level below which birthweight-specific mortality significantly increased. Low birthweight is further categorised into very LBW (VLBW, <1500 grams) and extremely LBW (ELBW, <1000 grams) (85). Birthweight is governed by two processes: gestational duration and fetal growth rate, as such LBW is the result of shortened gestation, impaired fetal growth, or a combination of both. The underlying causes of shortened gestation and impaired fetal growth are multifactorial, and their biological pathways differ. The LBW construct can be limited because it represents a mixture of infants with impaired growth, those with normal growth trajectory but delivered early and those small for genetic reasons unrelated to viability (constitutionally small). Consequently, this construct, although easy to measure, is not useful when the outcome of interest can be differentially related to shortened gestation and abnormal growth (76).

With advances in perinatal care services, the LBW cut-off is now rendered somewhat arbitrary however it is still a valuable indicator of maternal health, nutrition, healthcare delivery and poverty (77). LBW is particularly useful for clinical management in terms of indication for additional postnatal clinical surveillance and intervention; especially since LBW may be associated with long-term neurologic disability, impaired language development and academic achievement, as well as increased risk of chronic diseases (77).

2.2.2.2 High birthweight

High birthweight (HBW), also termed macrosomia, describes a newborn with an excessive birthweight. Establishing a universally accepted definition has proved challenging, with HBW defined as a birthweight of equal to or exceeding 4000, 4500 or 5000 grams, regardless of gestational age (66, 81). An HBW infant may be a result of maternal hyperglycaemia (either due to pre-existing or gestational diabetes), maternal obesity and/or excessive gestational weight gain, genetic or placental factors (87). While HBW is a condition of public health concern, this outcome has not received the same attention as PTD and LBW, particularly in LMICs even though it has been

associated with a range of maternal and perinatal complications (88-90). Infant complications include increased risk of birth asphyxia, traumatic injuries (linked to shoulder dystocia and brachial plexus injury) and perinatal mortality (89, 90). Maternal complications, which may be due to difficult labour and delivery, include increased risk of vaginal and perineal trauma, post-partum haemorrhage, infection, and hematoma (91-93). Given these complications, identification of pregnant women with macrosomic fetuses before vaginal delivery is attempted is necessary. This is however hampered by imprecise antenatal diagnosis with both clinical and ultrasonographic estimates of fetal weight particularly late in pregnancy when ultrasound is more unreliable (67, 94). More precise definitions of HBW predictive of morbidity or mortality are required in LMICs because there is an increased prevalence of maternal diabetes and obesity which could lead to increased prevalence of HBW infants (95, 96). This raises concern because in resource-limited settings with limited obstetric care services, pregnancy and delivery complications related to HBW could lead to severe adverse outcomes (88). Additionally, HBW is linked with increased risk of metabolic conditions such as hypertension in childhood (97) and increased risk of developing obesity and type 2 diabetes later in life (98).

2.2.2.3 Small for gestational age

Small for gestational age fetuses/infants are defined as having a weight or length below the 10th percentile for their gestational age, or 2 standard deviations below the mean (99, 100). In addition, weight or length less than the 5th or 3rd percentile or 2.5 standard deviation below the mean is used to denote severe SGA (101). The SGA outcome is commonly used as a proxy for growth restriction, especially in settings where serial ultrasonography is not readily available (102). However, despite many pathophysiological events being present that may result in growth restriction, SGA and growth restriction are not synonymous (103). This is because SGA only identifies small infants, however, appropriate for gestational age (AGA) fetuses/infants can also be growth restricted – since factors associated with fetal growth restriction affect the entire distribution (1). Using the terms interchangeably can lead to an overestimation of growth restriction among SGA and an underestimation or failure to detect growth restriction among AGA (53). For this reason, SGA is categorised into groups based on its aetiology: SGA due to being constitutionally small or SGA due to weight or length lower than expected optimal (104). When constitutionally small this is due to parental factors (e.g. maternal height, weight, ethnicity, parity) so weight/length are less than the 10th percentile however this is expected and there is likely no increased risk of perinatal mortality and morbidity (100). In contrast, SGA due to abnormal pathological processes are most commonly due to maternal (chronic conditions, infections, nutritional status, and substance use) or placental (placental haemangiomas, placental previa and chronic placental abruption) (104). These growth-restricted fetuses/infants are at increased risk of respiratory complications, hypotension,

hypoglycaemia, poor feeding, and hypothermia, which puts them at increased mortality risk (105). Later in life, SGA infants are at increased risk of delayed neurodevelopment (106), metabolic complications, and impaired reproductive functioning (107). The majority of SGA infants (approximately 70%–90%) show catch-up growth (108, 109). Failure to achieve appropriate catch-up growth results in persistent short stature as well as increased health risks and psychosocial impairment (110, 111). In order to distinguish between constitutional and pathologic SGA, individualised or customised growth standards should be prioritised because they set up optimal fetal and infant growth rates for each pregnancy, based on maternal characteristics such as maternal weight and height (112).

2.2.2.4 Large for gestational age

The terms LGA and macrosomia, which are often used interchangeably, both represent excessive growth however they differ slightly in their specific definitions (113). Macrosomia is defined as an absolute weight above a specified threshold regardless of gestational age while LGA fetuses/infants are defined as having a weight or length greater than the 90th percentile for the gestational age or 2 standard deviations above the mean (114). Consequently, LGA fetuses/infants could have weights less than 4000 grams. It has however been suggested that a definition of greater than the 97th percentile more accurately identifies fetuses/infants at the greatest perinatal morbidity and mortality risk (115). The clinical estimation of fetal size is difficult, so LGA is often not detected during pregnancy and labour (116). This has been highlighted by a comparison of EFW with actual birthweight which showed the poor sensitivity and specificity of ultrasound for correct prediction of an LGA pregnancy (117).

Similar to SGA, LGA is also categorised into groups based on its aetiology: LGA due to being constitutionally large or LGA due to weight or length higher than expected optimal because of a pathologic process (81). Fetuses/infants that are LGA due to parental factors tend to be healthy and likely will not experience adverse effects of their nutritional status (53). The most common reason for LGA fetuses/infants due to pathologic processes is maternal diabetes, however maternal obesity and/or excessive maternal weight gain during pregnancy are also contributory factors (118). Additionally, other factors such as genetics, intrauterine environment, nutritional status, and placental function also play a part (119). Being LGA, even if not always >4000 grams, is associated with an increased risk of adverse maternal and neonatal outcomes, with risks significantly increased as the percentile increases (120). Maternal risks include need to induce labour, caesarean section, and postpartum haemorrhage; while neonate risks include asphyxia shoulder dystocia, hypoglycaemia, as well as longer hospitalisation (120, 121). Additionally, long-term consequences of being LGA include cardiovascular conditions such as obesity, hypertension, obesity, metabolic

syndrome, and type 2 diabetes later in life (119-123). The growth of LGA infants is thought to slow down in infancy (124). However, after an initial period of compensation through a decrease in growth, the growth surfeit is expected to persist with LGA infants remaining around the 70th percentile for size through early childhood (125).

2.2.3 Pregnancy Loss

Pregnancy loss is defined as a spontaneous end to pregnancy without a live birth occurring between conception and the end of labour (1). However, since there is no marker for the occurrence of conception, this window could be narrowed to encompass the heterogeneous and broad spectrum of losses from implantation to delivery. Pregnancy loss can be identified either clinically (by bleeding or embryo/fetus expulsion) or ultrasound. A pregnancy loss that results in the passage of tissue can have confirmation of a diagnosis and identification of possible causes through a pathological examination (126). In contrast, if there is vaginal bleeding but no passage of tissue then an ultrasound examination will be required for further evaluation (127, 128).

Accurate classification of pregnancy loss is hampered by the lack of clarity in the nomenclature used. Traditionally pregnancy losses occurring before 20 weeks were classified as spontaneous abortions, while the losses after this period were classified as miscarriages and stillbirths. These classifications are however arbitrary, outdated, and not clinically useful because there is lumping together of losses which have different aetiologies (4). For instance, this classification separates a 'spontaneous abortion' at 19 weeks from a 'stillbirth' at 21 weeks (definition in some HICs) even though their aetiologies are likely to be more similar, while in contrast aetiologies of spontaneous abortions at 6 weeks and 19 weeks which are lumped together are likely to be different (126). Additionally, despite widespread ultrasound use for accurate clinical assessment and diagnosis of these losses, there are poor descriptive terms, such as missed abortion and blighted ovum, from historical classification systems which are still used (126). Use of developmental periods in pregnancy is more useful for classification because these are distinguishable (4, 129). In research studies, the combination of sensitive hCG measurements in serum and urine and ultrasonography has improved knowledge about early pregnancy development, enabling clearer descriptions of the gestational age at which a pregnancy loss occurs (130). The spectrum of pregnancy loss is subdivided into categories defined according to the gestational age at which they occur (Figure 2.2), consequently determination of accurate gestational age is essential.

Human reproduction has been described as an inefficient process with the loss of approximately half of all conceptuses before the expected menses, with an additional loss of 30% after or around the time of a missed menses (131-133). The vast majority of pregnancy losses (75%) are not from clinically recognised pregnancies because they are due to implantation failures (133). After implantation the pregnancy loss rate decreases to about 15-20% and then between 10-13 weeks the pregnancy loss rate is approximately 3% (126). An alternative view of human reproduction is that it is a very selective system, which is intended to optimise the final outcome of pregnancy (134); and so pregnancy losses have been thought to be a result of the selection out of fetuses with genetic or structural malformations. This is however not always the case with some losses being a result of maternal or environmental factors such as previous pregnancy losses (135). Correct classification of pregnancy losses enables better consideration of aetiology. Induced abortions and medical terminations of pregnancy are generally not considered as pregnancy losses (5).

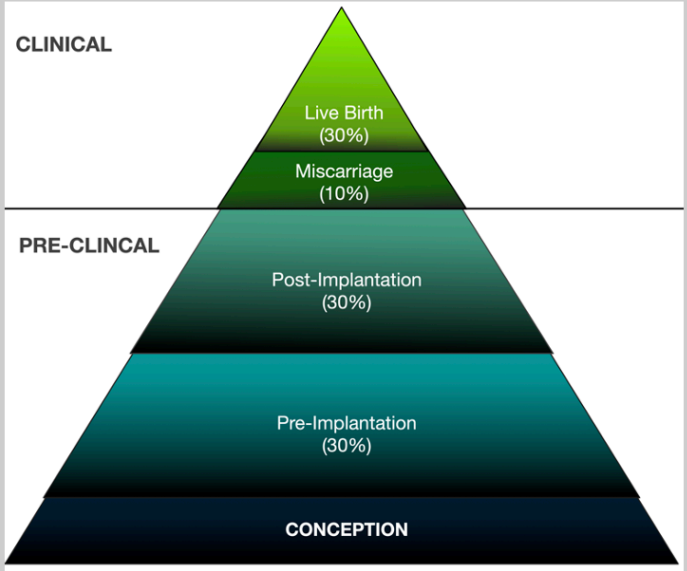

Figure 2.2 Pregnancy loss iceberg: overview of the outcomes of conception

Adapted from: Macklon NS, Geraedts JP, Fauser BC. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. *Hum Reprod Update*.2002;8(4):333-343.

2.2.3.1 Early Pregnancy Loss

Organogenesis is complete at the end of the ninth gestational week, heralding the end of the embryonic period and the start of the fetal period. Consequently, early (or subclinical) loss (EPL) is defined as a loss within 10 weeks of gestation (136). This type of pregnancy loss was previously classified as a spontaneous abortion. Because of the timing of EPL, an embryo may or may not be viable (have a reasonable chance of survival) prior to the loss depending on the developmental stage at which the loss occurred.

EPL can be categorised as either a pregnancy of unknown location (PUL), a peri-implantation loss, a pre-embryonic loss, or an embryonic loss (4). A PUL is defined as a pregnancy that is confirmed by a test, but a transvaginal ultrasound shows no signs of an intrauterine or extrauterine pregnancy i.e. pregnancy is not in the uterus (137, 138). Ultrasonography is the best examination method for identifying the location of an early pregnancy, and classifying a PUL as a definite ectopic pregnancy,

a probable ectopic pregnancy, PUL (with no ultrasound image of an intrauterine or ectopic pregnancy), probable intrauterine pregnancy or definite intrauterine pregnancy (138). Late diagnosis of an ectopic pregnancy is concerning because it may increase maternal morbidity and mortality and surgical interventions which may be required could negatively affect future fertility (139). The next type of EPL is a peri-implantation loss which occurs prior to 5 weeks gestation during the time leading up to and involving implantation and is detected by the absence of a visible gestational sac (126, 133). Following this, a pre-embryonic loss occurs during the 5th week of gestation and is detected by the presence of a visible gestational or yolk sac prior to the loss but no visible embryonic structure (4, 126). Lastly, an embryonic loss occurs between the start of the 6th week of gestation and the end of the 9th week of gestation and is detected by the presence of a visible embryo prior to the loss (126).

EPL is a very common outcome, with approximately two thirds of all pregnancy losses peri-implantation losses that often are not recognised as a pregnancy (133) (Figure 2.2). The majority of the remaining one third of pregnancy losses occur before 10 weeks of gestation.

2.2.3.2 Miscarriage

Miscarriage is defined as the spontaneous loss of an intrauterine pregnancy that is clinically recognised before a gestational age at which the fetus could survive outside the uterus (viability) (140). Challenges exist over the definitions of an unequivocal intrauterine pregnancy and the limit of viability, which either be defined according to gestational age or by fetal weight (141, 142). Historically, the miscarriage definition was a clinical loss before 28 weeks' gestation, because of the inability of fetuses to survive earlier than this. Depending on country and setting, the upper bound of miscarriages has now shifted to between 20 – 24 weeks (142). The frequent inclusion of the fetal weight criteria (<500 or <1000 grams) to the miscarriage definition is because of concerns around gestational age imprecision.

Three-quarters of pregnancy losses occur in the first trimester – this includes a combination of early pregnancy losses and early miscarriages which are frequently combined into one group. Miscarriages which occur during the fetal phase (>10 weeks' gestation) are also referred to as fetal deaths. Miscarriages are considerably less common than EPLs, with the frequency of losses decreasing with increasing gestational age (143). Miscarriages can be subdivided into those occurring early (10 - 15 weeks) and late (16 weeks – upper miscarriage limit in that setting) (4). While the date of a miscarriage is usually determined by the onset of bleeding and products of conception expulsion, this does not necessarily correspond with the date of fetal death which could have been much earlier. Clinical features of miscarriages vary according to when in pregnancy the

loss occurs, however many miscarriages can be detected by ultrasound before expulsion occurs (1). Miscarriages occurring early in pregnancy do not require medical attention as the bleeding is like menstrual bleeding, with the occasional passage of clotted tissues. Ultrasound examination of such losses show either an empty amniotic sac or a very disorganised embryo. In contrast, miscarriages occurring later in pregnancy have heavy bleeding and passage of recognisable fetal tissue and may require hospitalisation, sometimes with surgery to remove retained tissue (1). A definite diagnosis can be made by ultrasound if either of the following criteria is present: 1) passage of a conceptus with a CRL of at least 30 mm, 2) documentation of a dead conceptus with a CRL of at least 30 mm, 3) conceptus loss following documented fetal cardiac activity or dead fetus at delivery, with Apgar scores of 0 at 1 and 5 min (4). When the ultrasound is indeterminate and symptoms of miscarriage are present then serial hCG assessment could be used for diagnosis, although this is usually only available in research studies (144).

Miscarriage is a heterogeneous condition and the exact pathophysiology resulting in the uterine expulsion of early pregnancy remains unknown. Sporadic miscarriage, the most common complication of early pregnancy, primarily represent the failure of progression to viability by abnormal embryos. Fetal chromosomal abnormalities account for approximately 50% of all sporadic miscarriages, the other contributory factors include uterine malformations and/or hormonal problems (143, 145, 146). Recurrent miscarriage, a more severe miscarriage phenotype, is a less common phenomenon. It defined as two or three consecutive miscarriages (147), with previous definitions including three or more consecutive miscarriages (148). Recurrent miscarriage is considered as a distinct disease entity, with multiple aetiologies, which include maternal and paternal chromosomal anomalies, immune dysfunction, maternal thrombophilic disorders and various endocrine disturbances (149).

2.2.3.3 Stillbirth

Stillbirth is defined as the death of a fetus mature enough to have survived outside the uterus. However, the conversion of this definition to specific criteria has proved difficult, with stillbirths being the health outcome with the largest number of conflicting and legally mandated definitions (150). WHO previously recommended birthweight to be used as the first line determinant for stillbirth (151), however, gestational age has been shown to be a better measure as it gives an indication of developmental status (152). The gestational age cut-off point between miscarriages and stillbirths depends on legal and administrative definitions of vital events (livebirth and still birth) in different settings (142). Historically the gestational age criterion for stillbirths was a pregnancy loss at 28 weeks or greater. However, with improvements in medical care, survival outside the uterus at earlier ages has become possible, consequently the criterion has also changed. Multiple

definitions are used based on gestational age (minimum gestational age ranging from 20 to 28 weeks), birthweight (300 grams, 500 grams or 1000 grams), and/or an Apgar score (condition of the newborn in the first critical minutes of life) of 0 (153). The gestational age cut-off point tends to be higher in LMICs than in HICs, so depending on setting there could overlap with miscarriage. Given this variability, for the purposes of international comparability, WHO defines stillbirths as all infants born with no sign of life weighing ≥1000 grams, or after 28 weeks completed weeks of gestation (154, 155).

Stillbirths can be divided into three categories based on gestational age at delivery, into early (20 – 28 weeks), late (28 - 36 weeks) and term (≥37 weeks, particularly in settings where there are lower limits of viability (156, 157). This classification, while arbitrary given advances in neonatal care, enables the distinction between stillbirths that could potentially be prevented (late losses) and those that would be difficult to prevent (early losses) (156). Stillbirths can also be classified based on timing of fetal death either by whether death occurs before the onset of labour (antepartum) or during labour (intrapartum) (157). This roughly corresponds with the old concept of defining stillbirths as macerated or fresh, which helped determine timing of stillbirths before the advent of ultrasound and fetal heart rate monitoring (158). In LMIC approximately 50% of stillbirths occur during the intrapartum period and this is linked to inadequate obstetric care, while in HIC they usually occur in the antepartum period (159).

The aetiology of stillbirths depends on the time at which the fetal death occurred. Early stillbirths tend to be linked to infections, placental abruption, and fetal anomalies (160). Infections lead to stillbirths through several mechanisms: 1) severe maternal illness, 2) placental infection leading to oxygen and nutrient transfer disruption, 3) fetal infection leading to damaged vital organs or severe congenital abnormalities incompatible with life 4) precipitating preterm labour, leading to intrapartum death (161). Late and term stillbirths are frequently classified as unexplained, including those associated with growth restriction and placental abruption (162). Many unexplained stillbirths can be attributed to placental dysfunction, through either fetal growth restriction (an indication of poor placental functioning) or placental pathological changes. Fetal growth restriction is strongly linked to unexplained stillbirths, with approximately half of these stillbirths having a birthweight less than the tenth percentile (163). Histopathological examination of placenta has shown a relationship between maternal vascular malperfusion (MVM), a constellation of gross and microscopic findings representing abnormal maternal vascular channels perfusion (164), and stillbirths (165-167). Placental MVM which leads to poor placental development and abnormal blood supply, has also been associated with stillbirths in the presence of maternal hypertension and placental

haemorrhage (168). Fetal vascular malperfusion has also been implicated in stillbirths, although this is to a lesser extent than MVM (169).

2.3 Epidemiologic considerations for investigating birth outcomes

Perinatal epidemiology aims to investigate of the impact of factors inherent to the pregnant woman, with a goal of determining the causal contribution of these factors to the outcomes of interest (170). Events occurring during pregnancy that can influence outcomes include voluntary exposures, environmental exposures, diet, genetic constitution, the effects of infections and medical conditions, and medication use (171). The multiplicity of outcomes in this field introduces complexity because an understanding of each outcome's pathophysiology and the factors that affect each one is required. Factors affecting pregnancy outcomes are inter-related (Figure 2.3), and understanding this interplay is essential when conducting pregnancy-related research (171, 172).

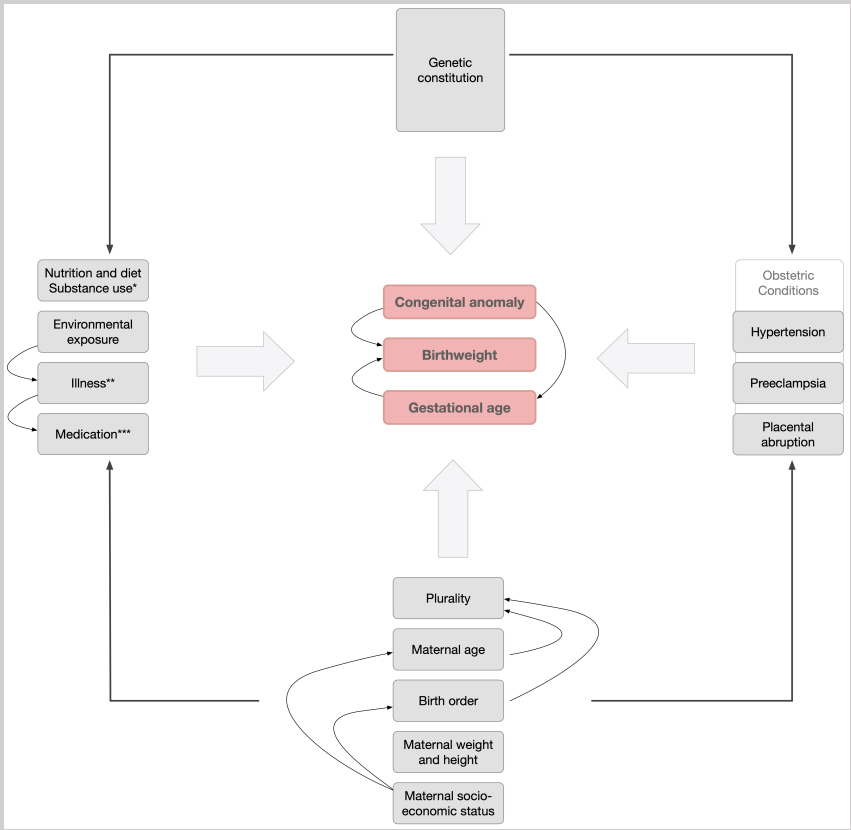

Figure 2.3 Inter-relationship of factors affecting the length of gestation, birthweight, fetal growth, and congenital anomalies

* Substances – alcohol, drugs, cigarettes
** Illness – an experience of ill health which may or may not be related to disease
*** Medication for illness, or other acute and/or chronic conditions

Adapted from: Petitti DB. Perinatal epidemiology: studying the effects of illness and medications during pregnancy. *Immunol Allergy Clin N Am*. 2000; 20:673-685.

In order to identify beneficial approaches for prevention or appropriate interventions for adverse outcomes, addressing causality through aetiological research is necessary. This needs to be done while fully recognising the challenges associated with accurate causal inferences. This requires discerning and understanding of not just patterns of associations between predictors and adverse outcomes, but also causal antecedents (172). The study of exposures and pregnancy-related outcomes is complicated by a number of methodological nuances, which underlie perinatal epidemiology specifically linked to the complexity of human reproduction function (1, 6, 173, 174). These nuances complicate the design and the interpretation of these studies, and inappropriate study conduct can have major implications for the correctness of its conclusions. Consequently, appropriate study design choice including suitable comparison groups, analytic decisions, and adequate reporting and interpretation of the results is vital.

When conceptualising perinatal research, a broader view of health is needed because with pregnancy, unlike most acute and chronic conditions, the progression and time course is predictable, with a key definable outcome to the health state – delivery of the infant (172, 175). Collection of data related to some outcomes of interest is problematic since many reproductive events are not illnesses, consequently systematic engagement with healthcare systems is not always required. Differing levels of access to outcome data means there is a wide range in the data quality ranging from near perfect to non-existent. Another challenge is that in the period from initial attempts to become pregnant until infant delivery there is an interrelation of events which involves complex selection phenomena and competing risks which could be easily mistaken for causal inference (1). Related to this reproductive outcome can be linked by common causes creating associations that link outwardly unrelated outcomes, if a condition such as subfertility leads to an increased risk of preeclampsia (176), treatment for this prior condition can be associated with a subsequent outcome. There are also notable analytical nuances and challenges to be considered. Identification of appropriate denominators i.e. the population at risk is a *sine qua non* of epidemiology, however some perinatal outcomes lack a true observable denominator (1). For example, conception cohorts are the appropriate denominators for many outcomes related to pregnancy loss, however only proxy denominators such as hCG detected pregnancies or women presenting for antenatal care are available (177).

The temporal nature of the reproductive process means that in both study design and data analysis careful attention needs to be paid to the timing of events. The time that a pregnancy comes under observation is important. This could either be when a woman first becomes aware of her pregnancy and thus potentially aware of a loss; or when the appropriate health care systems are aware of the pregnancy (and thus under observation in studies of women in antenatal care) (6). Explicit

recognition is required for the time at which the outcome and any exposures occurs. For some outcomes the timing could be the gestational age at delivery while for others the timing could be ambiguous and not observable at the time they are occur. Consideration of exposure timing is also important because while the duration of pregnancy is relatively short, exposures can be time-varying and change over the course of pregnancy (6). Given how specific the timing of embryonic and fetal events is, there is likely to be variability in the biologic consequences of exposures that change over time. The consequences of an exposure at one time in pregnancy could differ from exposures occurring several weeks earlier or later (178). Linked to exposure timing is dosage which could lead to different outcomes for instance a low dose could increase the risk of congenital anomalies while higher doses could increase the risk of fetal loss (179).

These are a few of the many methodological issues that apply to perinatal epidemiology, which require careful consideration. In addition to these, each of the outcomes have their own unique challenges linked to study design choices, exposure and outcome assessment and potential threats to validity that also require consideration.

2.3.1 Gestational Duration

2.3.1.1 Study Design Choices

Studies investigating gestational duration are typically focused on outcomes such as PTD and thus require accurate gestational age data for correct outcome classification. Therefore, the most appropriate study design for this outcome is a prospective cohort study, with recruitment as early as possible in pregnancy. Use of this study design enables the assessment of numerous exposures and risk factors and enables greater analytical flexibility. Other study designs can be employed when there are specific or unique types of outcomes or exposures being explored. For example, case-control studies can be useful when rare forms of perinatal outcomes are of interest such as the study of perinatal risk factors for necrotising enterocolitis in very preterm infants (180). Modified matched case-control study known as case-crossover can also be used investigating whether temporal variation in exposures, can be acute triggers of perinatal outcomes (181). An example is a study of transient factors (such as heavy physical exertion, acute infections, and stressful events) during the 72-hour period before the index time that could potentially act as triggers of spontaneous preterm labour and premature rupture of membranes (182).

Another study design that can be used is a time-series analysis, which is useful when the exposure (typically environmental) has short-term variations which are expected to have short-term effects (183). An example of such a study was one investigating the association between high apparent

temperature and PTD (184). This design enables elimination of potential confounding by individual risk factors that do not change over short time periods.

2.3.1.2 Outcome Assessment

While designing and conducting research in which gestational duration is the outcome, or related to the outcome of interest, decisions around which gestational assessment methods are used are vital. Fertilisation is a silent event, so there are substantial challenges in accurately estimating the conception date and by extension gestational duration. This can be defined either from LMP, typically collected retrospectively, either during pregnancy or after delivery; or from early pregnancy ultrasound measures, when fetal biometry varies linearly so CRL and BPD can be used as a proxy for gestational duration.

The differing levels of accuracy of gestational age assessment methods mean that studies using outcomes related to gestational duration, should whenever possible, collect gestational age data from different sources. This will require consideration of standard practises used. For instance, gestational age recorded in clinical records is expressed as completed weeks (rounded to the closest lower integer) rather than ordinal weeks – so for instance week 40 of pregnancy starts on day 280 which is the beginning of the 41st week. This clinical approach results in considerable loss of information compared to ultrasound measurement where data on weeks and days is available. Using ultrasound-determined gestational age allows for gestational duration to be treated as a more continuous measure such as days. This is preferable particularly if gestational duration is to be analysed as a continuous outcome or as an adjustment factor in the study of birthweight (5).
With the increase in obstetric interventions through inductions and caesarean sections, which directly impact gestational duration, detailed information on the onset of labour is also important. This will enable distinction between spontaneous and medically indicated delivery.

2.3.1.3 Potential Threats to Validity

The main source of bias in studies of gestational duration is related to the choice of assessment method. The menstrual dating approach suffers from two distinct problems with LMP-based estimates leading to errors in either direction. First, erroneous recall leads to random errors which is reflected most explicitly in digit preference in the recall LMP dates. This manifests as an over-reporting of days rounded to the nearest 5 (185); and the more problematic over-reporting of the 1st, 15th and 28th an indication of rounding to the beginning, middle and end of the month, incurring substantially more error (186, 187). In contrast to this random error, there is a clear directional error superimposed that is reflective of inaccuracies due to mid cycle bleeding and/or bleeding early in pregnancy occurring after LMP date, making gestational duration seem shorter than it actually is

(17). Conversely long menstrual cycles and delayed ovulation falsely lengthen the apparent duration, which appears to account for much of what is classified as post-term deliveries (188). Bias can occur if error in LMP-based estimates of gestational age is differential with respect to exposure, for instance if the exposure of interest influences the duration of menstrual cycle duration or its regularity (6).

The ultrasound approach relies on the assumption that fetal growth is similar for all fetuses during the first trimester, which is why early ultrasound estimates are considered the most accurate measure of gestational duration (187). However, measurement error of fetal measures is possible and since any natural variations in fetal growth rate are automatically translated into fetal age variations this could introduce bias. Additionally, systematic errors in fetal measures (and thereby gestational age) could be due to the calibration of the ultrasound equipment or the choice of fetal measure-gestational age standard both of which can affect estimates. For most pregnancies, these errors may be of minor significance and the misclassification will not affect comparisons of the effect of different aetiological factors on gestational age. However, in some cases this measurement error may complicate the interpretation of fetal risk associations when the error is associated with the aetiological factor in question. For instance, an exposure that interrupts fetal growth from as early as the first trimester but has no effect on the gestational duration could lead to systematic bias being introduced because the smaller fetuses are assigned a lower gestational age than the larger normal fetuses even if in actuality, they are the same gestational age (187). In these cases, the less accurate but also less prone to bias measure based on LMP could be preferred (187). Selection bias is another potential problem associated with the use of ultrasound when investigating associations with perinatal outcomes – if ultrasound is only available to a proportion of population the benefit of ultrasound adjustment will not apply equally across the population, which would potentially confound the epidemiological analyses of gestational age-related outcomes.

The other methods of assessment are also susceptible to errors that can introduce bias. Measurement of the SFH suffers from poor reproducibility and its accuracy can be diminished by a multiple pregnancy, high body mass index (BMI), IUGR and other maternal or fetal characteristics (189, 190). There has been a proliferation of postnatal assessments, however problems still exist with their accuracy and implementation (191), and data suggests that these methods are the least accurate – resulting in gestational age overestimation in infants born before 40 weeks GA and GA underestimation in infants born at or after 40 weeks (192, 193).

In addition to issues related to gestational age assessment, bias can also be introduced during the analysis phase. For instance, associations in highly selected populations, such as analyses restricted

to preterm deliveries, can complicate interpretations. In such cases, only certain aetiological questions can be explored without incurring selection bias, because while estimates of the association between a given exposure and outcome in the restricted population may have prognostic value, they do not necessarily reflect causal effects (194). In assessing the association of exposures with a birth outcome, it has long been routine to adjust for gestational age, even though adjustment for an intermediate variable can produce bias when unmeasured confounders act on both the intermediate and the outcome (195). When gestation-adjusted associations are calculated in the presence of uncontrolled mediator-outcome confounding, collider-stratification bias is introduced (196, 197). There are a few scenarios in which gestational age itself can be a confounder – gestational age would have to be a cause of the exposure or outcome of interest (198). Outside of these scenarios, refraining from adjustment is the only sure way to avoid potentially damaging collider bias (18).

2.3.2 Fetal Growth

2.3.2.1 Study Design Choices

Birthweight is one of the most accessible epidemiological variables since it is usually precisely recorded and widely available. As with most outcomes, prospective cohorts with antenatal recruitment represents the preferred study design option, particularly when exposures with toxicological relevant windows are being assessed. If the outcome of interest is fetal growth rate or restriction/overgrowth, then recruitment as early as possible in pregnancy is desirable so that fetal size evaluation can take place given that the gestational and yolk sacs are visible from as early as 4 to 6 weeks gestation. Unlike other outcomes, birthweight information from clinical records and from self-report is of usually of high quality (especially if collected/recorded close to delivery); so, study designs employing retrospective data collection are also possible because outcome misclassification is limited. Important to note however is that using retrospective designs makes the studies subject to exposure and confounder misclassification, importantly, from heterogeneity and error in gestational duration assessment.

The distinction between fetal size and fetal growth is important because different patterns of fetal growth and exposures may result in the same neonatal anthropometric measurement. Consequently, a longitudinal study of fetal growth using repeated ultrasound measures would enable identification of specific critical periods for the impact of the exposure on fetal growth (199). Using this longitudinal approach limits the delay between exposures and fetal growth measurements, reducing potential confounding and makes use of exposures concurrent with the observed fetal growth pattern (200). In addition to prospectively collected longitudinal data, analytical methods accounting for repeated measurement correlation on a single fetus would need

to be employed (201). An example of a study using this approach investigated the association between ambient air pollution and fetal biometry during the second trimester (200).

2.3.2.2 Outcome Assessment

While birthweight is a reliable measure, it is not without error or variation depending on the timing of measurement. Consequently, in order to get an accurate measure of birthweight, this should ideally be measured within first hours after birth, since birthweight can be affected by meconium, urine passage or timing of any feeding relative to weighing which leads to a loss of 3-10% of infant weight between 24 to 72 hours after birth (5). Birthweight is a summary measure that does not identify the respective contributions of the different organs. To address this limitation, in some studies other measures of neonatal anthropometry (length, head and abdominal circumference) are also taken to provide additional information. These anthropometric measurements are also affected by timing, however unlike birthweight, these measures should be measured from 24 to 48 hours after birth to enable correction of labour-related distortion. For instance, birth length measurement can be biased by the infants flexed state, which takes several days to resolve and as the degree of flexion decreases the mean measured length can increase by ~0.2 cm within the first 24 hours and by a further 0.2 cm by 48 hours (17).

By using projected growth trajectory through the course of pregnancy, possible periods of fetal growth retardation can be identified and quantified through deviations in actual growth relative to the projected pattern (202-204). Assessment of fetal growth by ultrasound can also be useful for the detection of possible transitory exposures effects on fetal growth; and accurate identification of IUGR could potentially elucidate specific mechanisms related to exposure that led to growth. In studies using ultrasound measurements, consideration should be given to also collect details on personnel performing the measurement, and if possible, limiting the number of sonographers and devices used. When not possible, standardisation workshops should be conducted prior to the study.

2.3.2.3 Potential Threats to Validity

The main source of bias in studies that use size at birth as a proxy for assessing fetal growth, is that this measure does not adequately consider the gestational time points during which growth is slowed or affected, and consequently a late assessment of growth restriction could potentially introduce bias. Firstly, the common practice of classifying infants at birth as being growth-restricted just because they are small for gestational age could bias effect estimates since these two phenomena are not synonymous (205). Secondly, assessing fetal growth at the end of the intrauterine period poorly reflects growth restriction because it masks growth patterns throughout

fetal life (in particular in the first two trimesters) and lacks consideration of critical periods of growth during gestation (206). This also ignores the possibility that while growth restriction affecting long term outcomes could occur at some time, fetal growth could continue and by birth achieve population growth standards i.e. various growth trajectories could result in the same birthweight (207, 208). Additionally, low birthweight can result from both growth restriction and shortened gestational duration (94), so determining which factor or combination of factors contributed to the birthweight is difficult when weight and anthropometric measures are only assessed at birth.

Studies with recruitment during the postnatal period are subject to potential confounders misclassification since these will be measured/assessed after pregnancy when they should have been assessed during pregnancy. An example of such an exposure is maternal smoking which is best assessed during pregnancy (5).

2.3.3 Pregnancy Loss

2.3.3.1 Study Design Choices

The study design chosen to investigate pregnancy loss depends on the research focus (category of loss), with different options depending on whether it is an early fetal loss, miscarriage, or stillbirth. Prospective cohorts with recruitment preconception or early in pregnancy are the most appropriate design for establishing risks of pregnancy losses. When pregnancy ascertainment and study recruitment only occur when medical attention is sought, women in these studies enter observation at different gestational ages and the earlier this observation commences increases the likelihood of fetal loss detection (209).

For early pregnancy loss questionnaires-based approaches are not ideal because of the select nature of the group of women in whom would be available, consequently under ascertainment would be a major concern with small proportions of losses identified. If the outcome of interest is miscarriage, recruitment of the majority of women should ideally occur before the end of the first trimester (210). This provides the most accurate estimate of miscarriage risk. On the other hand, if the outcome of interest is stillbirth, although early recruitment is desirable any recruitment before the end of the second trimester could be sufficient (210). For both these outcomes, while prospective data is considered to be gold standard, this data is sensitive to the entry time and length of observation time of the pregnancies, with early detection leading to longer observation time and consequently the higher the cumulative risk of loss (1). On the other hand, retrospectively collected data is subject to under-ascertainment of the earliest recognised losses, because there is a reliance on clinical records or self-report, both of which tend to be incomplete. In addition some losses (miscarriages) may be recognised by the mother but never recorded in clinical notes (1). For both

miscarriages and stillbirths, questionnaire-based approaches can be used since these losses usually occurs after pregnancy ascertainment and detection is either by the woman, sonographer, or clinician (5). Great care is however needed when designing these questionnaires to ensure that the outcomes can be distinguished.

Analytical considerations are also important when designing studies of pregnancy losses. Most studies of early pregnancy loss and miscarriage are left truncated because of antenatal recruitment after conception. Consequently, women are at different gestational ages at study entry and an unknown proportion of the conception cohort are missing because of pre-enrolment pregnancy losses (180). In order to account for these pre-enrolment losses, requires the use of survival analysis (211-213). This approach allows time under observation to be accounted for, through a series of risk sets with inclusion of only women who have entered the study by the time of each observed pregnancy loss (209, 214).

In cases of induced abortions or terminations of pregnancy, right censoring (another event occurs before an event of interest occurs, or the study ends before the event has occurred) is a possibility because these terminations may be precluded by a spontaneous abortion, consequently competing risk models are required (215).

Previous history of miscarriage is often predictive of the outcomes of subsequent pregnancies (211-213), and so reproductive history variables are typically included as explanatory variables in analytic model. While this approach may be appropriate for predictive models, when the goal is to understand aetiology, this should be avoided as it may induce bias (214, 216). This is because reproductive history does not always correspond to a confounder especially if the previous adverse outcome was the result of a persistent exposure corresponding to that under study (217).

2.3.3.2 Outcome Assessment

Assessment of early pregnancy losses is challenging because outside of in vitro fertilisation identification of timing of conception is difficult especially because at women are often unaware of their pregnancy. This is compounded by the fact that these losses do not have any accompanying clinical signs besides vaginal bleeding which could be mistaken for menstrual bleeding. Detection of early pregnancy losses would require measurement of biological markers, such as regular urine sampling, of women attempting to become pregnant. These samples will enable a precise detection of the post-conception rise in hCG levels (218), followed by the subsequent drop associated with embryonic or fetal loss (219). These types of studies are also appealing because they enable

concurrent exposure and outcome measurement thus circumventing the problems of retrospective measurement of fetal losses (220).

Miscarriages are usually self-detected or by a clinician or sonographer. Clinically, the occurrence of vaginal bleeding, abdominal pain and/or the loss of pregnancy symptoms are suggestive of miscarriage (128, 221). Speculum examinations are useful for diagnosing miscarriages when women present with heavy bleeding and signs of cardiovascular instability (128). This examination enables detection of placental and/or fetal tissues (retained products) that remain in the uterus after spontaneous pregnancy losses and facilitates their immediate removal. Pelvic examinations (including speculum investigation) for miscarriage diagnosis are unreliable, when clinically stable women present with mild to moderate vaginal bleeding in early pregnancy (128). In these cases, vaginal and speculum examinations have been shown to be largely inaccurate when compared with ultrasound findings (222).

Ultrasound diagnosis of complete miscarriage is made when there are no signs within the uterine cavity of any pregnancy tissue. This diagnosis can only be definitive if previous ultrasound examinations showed clear evidence of intrauterine pregnancy (221). In the absence of a previous scan, serial hCG measurements are required over 48 hours (223), with a decline by 13% or more indicative of a failing pregnancy (224). Other morphological features which are not diagnostic but are indications of increased miscarriage risk include presence of a yolk sac without a viable embryo, irregular gestational sac with a thin trophoblastic layer, size discrepancies between gestational sac and embryo, and bradycardia (221, 225, 226).

While a pregnant mother may suspect an *in utero* stillbirth based on absence or decreased fetal movement, this is not always indicative of a stillbirth (227). Likewise physical examinations are of limited value in the diagnosis of stillbirths, while inability to obtain fetal heart tones upon examination suggests fetal demise; however, this is not diagnostic, and death must be confirmed by ultrasonographic examination (228). Ultrasonography is essential for an accurate stillbirth diagnosis, with visualisation of fetal heart and demonstration of the absence of cardiac activity (227).

2.3.3.3 Potential Threats to Validity

The main sources of bias in pregnancy loss studies are related to the variations in gestational age at pregnancy recognition or at study recruitment/inclusion, which lead to different lengths of observation time at risk and the exclusion of some early losses. Because of these different observation times, a problem associated with these studies of pregnancy losses relates to left truncation and censoring issues.

In general, conception (time zero) is unmeasurable on a relevant timeline (gestational age) and more so in prospective pregnancy studies, where women are recruited at varying time intervals including after becoming pregnant or after having experienced unrecognised pregnancy losses. Consequently, because of pregnancy losses prior to enrolment an unknown proportion of the source population is missing, and women with pregnancies ending in livebirths are more likely to be in the study (214). This truncation is further complicated by interval censoring because of the considerable uncertainty underlying the exact timing of embryonic or fetal loss. It is important to be cognisant of left truncation and interval censoring when designing and analysing studies related to fetal loss because without accounting for these two phenomena bias may be introduced since study entry is differential by outcome and potentially also by exposure status (214). While less prevalent, right censoring or competing risks in the form of pregnancy terminations and other outcomes of pregnancy such as ectopic pregnancies, also need to be considered.

In addition to truncation and censoring, pregnancy loss is a challenging end point to ascertain because the exact timing of the loss is often unknown with the assigned timing usually based on the time of clinical recognition of pregnancy loss, which may occur several weeks after the actual death of the fetus. Because the different pregnancy loss outcomes are classified according to the gestational age at which they occurred, depending on gestational age ascertainment method used, there could gestational age measurement error, potentially leading to misclassification e.g. miscarriage classified as stillbirths. Additionally, there could be error in the estimation of gestation at the time of loss since this usually self-reported, sometimes months after the event.

2.4 Bias in birth outcomes

Human reproduction and development by nature is a highly timed and interrelated process underscoring the usefulness of prospective studies with longitudinal data collection across sensitive windows. In the hierarchy of study designs, evidence from randomised control trials (RCTs) is typically considered to be of the highest grade, because better control over possible bias is provided through randomisation and blinding. Theoretically, randomisation is ideal for drawing strong inferences about the effect of an exposure on maternal, fetal, and perinatal outcomes, since it ensures that the exposed and unexposed are exchangeable in terms of factors other than the one being studied investigated (171, 229). However, the involvement of pregnant women in clinical trials is a complex issue, with several obstacles existing in because randomising women is not always feasible for ethical and logistical reasons, particularly in the early stage's disease or outcome aetiology evaluation or when investigating the prognosis of short or long-term consequences of an exposure or condition. Even when an exposure can be randomly assigned, attaining an adequately

powered trial able to detect significant differences in clinical outcomes, in a reasonable amount of time, can be challenging (172, 230). Additionally, application of strict inclusion and exclusion criteria often lead to the inclusion of a very small and specific proportion of the population of interest, meaning the trial population does not reflect the 'real world' and results are not generalisable (231). Consequently, observational studies play a central role in aetiological and prediction research (232) and are becoming increasingly important for causal analysis in perinatal epidemiology given the practical and ethical costs of conducting randomised trials.

In order to assess causal relationships between exposures and outcomes, accurate measurement of exposure prevalence, disease occurrence and relevant covariates in representative study populations is essential. However, perinatal epidemiology studies like all epidemiological studies, in particular those of an observational nature, are invariably subject to bias. This is any systematic error that occurs during the design, implementation, or analysis phase of the study such as selection bias including cases of loss to follow-up, ascertainment and information bias. The consequence of bias is an incorrect estimated association between exposure and outcome of interest leading to deviations from the true measure of association (233). While random error can be reduced by increasing sample size, bias can only be reduced by changing the design and implementation of the study. In addition to implementing bias minimising strategies, unavoidable biases need to be identified and their potential impact assessed because of the implications this has for validity of the study.

This section will focus on more specific cases of information and selection bias often encountered in perinatal epidemiology and typically inadequately addressed. Because of the difficulty of assessing selection bias, this thesis will focus primarily on measurement bias and correction for it, while only briefly covering selection bias.

2.4.1 Information Bias

Information bias is an umbrella term that includes all biases that result in the inaccurate measurement or classification of the exposure status, outcome status or both (234). The most important type of information bias is measurement error or misclassification bias (235), with a certain amount of error is intrinsic to any measurement process. Other information bias subtypes exist which include observer/interviewer bias and recall bias (236).

The systematic measurement error or misclassification of participants on one or more variables related to exposure and/or outcome leads to distortion of association estimates between exposures

and outcomes. This can occur in a number of ways during study design, data collection or analysis (237). During the design phase this bias can arise conceptually, when there is discord between study definitions and true definitions which can lead to inaccurate operationalisation with respect to dose, duration, or induction periods (238). This definition discord can lead to some participants being incorrectly classified with respect to their aetiologically relevant exposure status. Bias can also occur during the conduct of the study when information is not accurately collected or measured through the use of faulty instruments, inaccurate measures or reporting and data entering errors (238). An example is assessment of gestational age which can be estimated using various methods that are all prone to some degree of mismeasurement, not all of which are at random (236). Single cross-sectional measurement of variables, such as blood pressure, that vary over time instead of multiple measurements is another example of inaccurate measurement. Lastly, during the analysis phase bias can be introduced through analytical decisions such as the incorrect categorisation of continuous data (237, 239).

2.4.1.1 Measurement error

Measurement error of exposure, outcome and covariate measures are widespread in epidemiological studies. In perinatal research, gestational age is often included as an outcome or covariate in many studies, however accurate measurement is difficult because all methods of assessment are subject to a degree of measurement error. Gestational age estimates are particularly prone to error they are based on routinely collected LMP, which can be unreliable because of inaccurate recall (240), intentional masking (241), and misinterpretation of early pregnancy bleeding as normal menses (242). Early ultrasound-based gestational age estimates are thought to be the most accurate however they could also be biased due to variations in fetal growth already evident by the second trimester of gestation (243).

It is a widely accepted notion that measurement error of important variables such as gestational age introduces bias and imprecision in exposure-outcome associations (244). However, despite this, disproportionately little attention is paid to the impact on estimates of effect. Several textbooks (245-248), methodological reviews (249, 250) and toolkits (251) have demonstrated examination, quantification, and correction of measurement error in epidemiology however most of this work has focused on error in exposure and confounding variables given their perceived greater impact on association relationships than measurement error in the outcome (245).

A fundamental prerequisite for analysing measurement error is specification of the models that quantify the relationship between the truth variable (unobserved) and the measured variable (observed). Types of measurement error are considered according to whether the measured

variables of interest are continuous or categorical because this determines how they will be dealt with.

Continuous Variables

For continuous variables, the common types of errors are distinguished by how the measured variable relates to the truth (244). The classical error model (Equation 1.1) is the model in which the observed variable is the truth measured with additive error; while the Berkson error model (Equation 1.2) is a reversal of this, with the truth the observed variable measured with error (246). In assessing measurement error careful attention must be given to the type and nature of error and the sources of data that allow modelling this error.

Box 2.1 Equations of classical and Berkson error models

$$\mathbf{X^* = X + U} \quad (1.1)$$

$$\mathbf{X = X^* + U} \quad (1.2)$$

X = Truth Measure

X* = Observed Measure

Within the classical error model, different types of error can occur. Random measurement is the simplest type and occurs when the measured variable can be expressed as the true variable plus a random component with zero mean and constant variance (246). The random nature of these errors means that multiple measurements will result in values above and below the true value. For example, with gestational age assessment by ultrasound, random errors are typically due to observer inconsistency in selection of measurement plane or placement of measurement callipers; and they can be reduced by averaging several measurements (252).

The other type of measurement error related to the classical error model is systematic error. This is when the measured variable is a biased representation of the true variable to such an extent that repeated measurements would not approach the true value (238, 246). Systematic errors are concerning because their consistency in direction and size (either consistently higher or lower than the true values) mean they can be difficult to detect; and they can cause bias in the exposure-outcome association in either direction (246). LMP and ultrasound-based gestational age measures for some subgroups of the population have been shown to be subject to systematic errors (243, 253). These types of error often result in somewhat large (>7 days) discrepancies between LMP and ultrasound-based measures, which can result in spurious associations between certain covariates and outcomes related to pregnancy length (187).

In contrast, with the Berkson error model, because the observed variable remains fixed in repeated sampling while the true variable varies, any error is independent of the true or measured exposure. Consequently, this type of error reduces power of the study by diminishing precision however causes little or no bias in measurement (246).

Categorical Variables

Measurement error in categorical variables is referred to as misclassification. If the true categorisation is known with respect to a measured variable (exposure or outcome), the likelihood of the classification system correctly grouping participants can be assessed (238). Four measures summarise the performance of the classification system – sensitivity (SE), specificity (SP), positive predictive value (PPV) and negative predictive value (NPV) (238). The probability that a participant that truly has the characteristic is correctly classified is the classification systems sensitivity. Likewise, the probability that a participant who truly does not have the characteristic is correct classified is the classification systems specificity (254). These two measures give an indication of the quality of the classification system. Once a participant has been classified, the probability that the assigned classification was correct is measured by the PPV and NPV.

In many cases, there is concern that the ability of the classification system to classify participants with respect to some variable may depend on the value of another variable and this concern can be described as non-differential and differential (238). Non-differential misclassification occurs when errors in exposure or outcome status classification occur with approximately equal frequency in the groups being compared. For instance, exposure status is determined before the occurrence of the outcome of interest or when method of assigning exposure is blind to outcome – errors in exposure status are not likely to be related to outcome status. Differential misclassification occurs when errors in exposure or outcome status classification occur differently in the groups to be compared. Differential exposure misclassification is likely to occur when the exposure is assessed after outcome occurrence – having the outcome may trigger more recall of perceived causes than in subjects without the outcome. Differential outcome misclassification is likely to occur if exposed participants were more likely to have their outcome ascertained than unexposed participants. When a binary exposure is related to an outcome, random non-differential misclassification present in the exposure will result in the attenuation of the exposure-outcome association (255). However, when a polychromous exposure is subject to systematic or differential misclassification, or when confounders measured with error are included in analysis models, prediction of the direction of bias in the estimate of the true exposure-outcome association bias is difficult (245).

2.4.1.2 Outcome measurement error/misclassification

Both outcome and covariate variables are equally likely to be mismeasured. However, measurement error in covariates has received extensive research interest, with a large body of analysis methods, whereas less attention has been paid to mismeasured outcomes. With a continuous outcome variable described by a linear regression model, if a linear form is assumed, measurement error increases variability of fitted line and decreases power for detecting effects however it does not cause bias (246). There is a profound and important difference when the outcome is binary and subject to misclassification. Whereas there is a near null impact of error in linear regression there is a major impact with outcome misclassification which can lead to major biases in parameter estimates thus needs to be accounted for (246). In addition, when measurement error exists in both covariates and outcome variables additional complexity is introduced with more complex effects on altering the model structure than measurement error in the outcome variable alone (254). This is of major concern in the study of birth outcomes because gestational age is typically categorised into outcomes such as preterm or post-term delivery, as reflective of gestations at which there is increased risk of neonatal mortality. Consequently, the implications of inaccurate gestational age could be greater with these outcome variables because of the potential for misclassification which could lead to invalid study inferences.

2.4.1.3 Potential solutions for measurement error

When assessing any measurement error for possible remedies, one central recommendation is to develop and to use more reliable measures. However, this is not always possible and other solutions are required such as correcting for the consequences of measurement error.

Even with improved measurement, some error will almost always remain in the data. Despite extensive methodological literature on how to mitigate the biases which are introduced to statistical estimates by measurement error (248, 256), these methods often not utilised. There are two general approaches which can be employed, and these are estimation-based or imputation-based approaches (257). The goal of the estimation approach is to mitigate coefficient and standard error bias within a statistical model; with the original dataset remaining unchanged with only corrections to the estimates (257). Consequently, for these methods the error structure and its mechanism in the analytical models needs to be known. In contrast, imputation approaches generate numerous separate datasets which would have been observed if there was no measurement error. Any statistical model can then be used to analyse the datasets. For this approach, in order to generate the imputed datasets, the error structure also needs to be known.

It should be noted that despite these differences, for both the estimation and imputation approaches approximate information is required about the error structure of the variables in question. The ideal scenario would involve direct measurement error estimation; however since this is not always possible even without empirical error structure information, sensitivity analyses with different combinations of reasonable best- and worst-case scenarios of measurement error are an option (257). This would still be more appropriate than an analysis based on the premise that no measurement error had occurred at all.

Estimation-based correction

Many methods have been described for correcting for the effects of measurement error in regression models (246). The most widely used method for correction of continuous variables is regression calibration because of its simplicity and applicability in different types of regression models. However, this method is used primarily for adjustment of measurement error in exposures. In regression calibration, the true exposure (unobserved) is replaced when fitting the outcome regression model by the expected value of the true exposure, conditional on the measured exposure and the other error-free covariates for each individual. This method gives consistent estimates of the true associations between the explanatory variables and the outcome in a linear regression model, and approximately consistent estimates in non-linear models, including logistic regression models (258) and Cox proportional hazards models (259).

For categorical variables, misclassification can be corrected using six measures of classification: SE, SP, PPV, NPV, false negative proportion and false positive proportion (238). Simple bias adjustment is conducted at the summary level and one fixed value is assigned to each bias parameter giving a single revised estimate of the association (260). Multidimensional bias adjustment which is also conducted at the summary level provides more information than simple bias analysis, with more than one value assigned to each bias parameter giving a range of revised estimates of association (238, 260). This method enables one to determine what combination of bias would be required to yield a null result. The limitations of these two methods are that they do not explicitly incorporate uncertainty around the bias parameters in interval estimates or tests of the target parameter and are unable to adjust for confounders. In contrast probabilistic adjustment which can be conducted at both the summary and record level, is able to incorporate uncertainty into statistical results by assigning probability distributions to each bias parameter giving a frequency distribution of revised estimates of association (238, 260).

Imputation-based correction

Multiple overimputation is an appealing alternative to estimation-based error-correction approaches (261, 262). Multiple overimputation extends the idea of multiple imputation of missing values to the mitigation of measurement error. Each observation of a variable is understood to be placed on a continuum reaching from perfect measurement, increasing grades of measurement error to no measurement at all. The multiple overimputation procedure creates multiple data sets by imputing all values conditional on observation-level priors and all other variables in the dataset with an expectation–maximization algorithm (257). The priors for the mismeasured variables can be estimated within the procedure from gold standard or proxy variables.

Multiple overimputation has some advantages over estimation-based methods. The procedure is, in comparison, quite robust to violations of the random normal error assumption and less dependent on the analytical and correction models (261). After the creation of multiple over-imputed data sets, any data transformations can be conducted, and any familiar statistical model can be used as is the case with "normal" datasets (262). An added benefit of this procedure is that it does not only take measurement error into account, but also deals with missing values.

2.4.2 Selection Bias

Selection bias, like confounding, is a mechanism that generates pathways between the exposure and outcome that are non-causal. The distortion this bias results in can be either towards or away from any true association, potentially masking a true association or a true lack of association (263). The existence of selection bias and any impact it might have been more difficult to recognise than confounding or measurement error (264). Selection bias in literature is primarily described in relation to where selection of observations into a sample is not independent of the outcome (265). However many other scenarios can create this bias so endogenous selection bias has been suggested as an umbrella term that includes sample selection and the other biases typically associated with selection bias (264, 265).

Endogenous selection bias can emanate from sources related to both study design and analysis. The first occurs at the beginning of a study when an unrepresentative study population is selected which does not reflect the underlying population of interest. This is of particular concern in perinatal epidemiology where study population, which range from preconception to birth stages, are often difficult to define. This is particularly challenging because the underlying populations are incompletely observed due to the high natural attrition that occurs from the preconception period through to birth (266).

Secondly, it can occur during study follow-up through informative censoring from differential loss to follow-up, which also results in an unrepresentative study population. This bias can also occur if a sub-set of participants are selected for more detailed investigation based on volunteering for further follow-up or having most data available (263). Lastly it can occur during the analysis stage by conditioning on a collider, or a variable influenced by the collider through to restriction (by design or analysis), stratification or regression adjustment (267). A collider (Z) is independently caused by two variables (X *and* Y*)* i.e. it is a common effect. Conditioning on Z will bias the estimated causal association of *X* on *Y because* this action is the same as observing the association in a subset of the population where all individuals have the same value of *Z* (Figure 2.4).

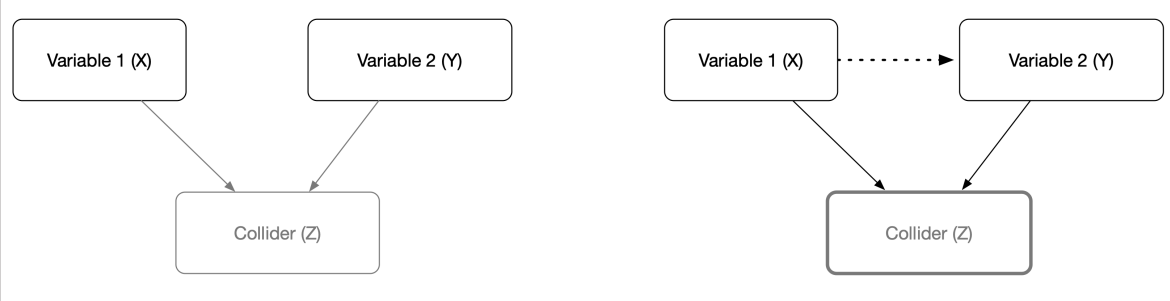

Figure 2.4 Illustration of selection bias through collider

When there is no relationship between X and Y, conditioning on Z can lead to selection bias through the opening of the backdoor path (**Variable 1 → Collider ← Variable 2**) which results in a spurious association between X and Y.

Adapted from: Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol.* 2018;47(1):226-235.

This bias can either be towards or away from the true association and can distort a true association or true lack of association (197, 263). This type of selection bias is of concern in perinatal epidemiology - particularly related to how gestational-age-specific or birth-weight-specific associations, that lie on the exposure-outcome causal pathway, should be dealt with (237, 268).

Endogenous selection bias can result in inaccurate estimation of relationships between variables thus affecting the internal validity of study. External validity of a study is also impacted because results from a biased sample may not be generalisable to other populations (269).

2.4.2.1 Types of selection bias

Processes of selection and attrition during the human reproduction process render underlying populations incompletely observable because gestations are removed from the "at risk" population (266). Consequently there are specific cases of selection bias linked to birth outcomes that are encountered (and typically inadequately addressed) in perinatal epidemiology, which will be highlighted here (268).

Live birth Bias

Over the gestational period extreme cohort attrition takes place with only 60–70% of all fertilised eggs likely to result in a live birth (133, 270). Under certain conditions this cohort attrition can induce selection bias because of the inability to account for the lost fetuses in risk estimation (268, 271). When long-term outcomes (infant and/or developmental), which can only be determined after birth, are of interest, only live born infants can be included in the at-risk population. This restriction to only live births results in live birth bias, a type of selection bias common in perinatal epidemiology that could lead to estimates of exposure on outcomes being biased (271, 272). There are two selection mechanisms that can result in live birth bias (271, 273), these are referred to as "depletion of susceptibles" and "conditioning on a collider" (274). These two mechanisms are both forms of collider-stratification bias where restriction to live births induces a spurious association between the exposure and the collider which in turn which results in a biased exposure-outcome association (275). Despite being parameterised differently these mechanisms are indistinguishable on a directed acyclic graph (DAG) (Figure 2.5).

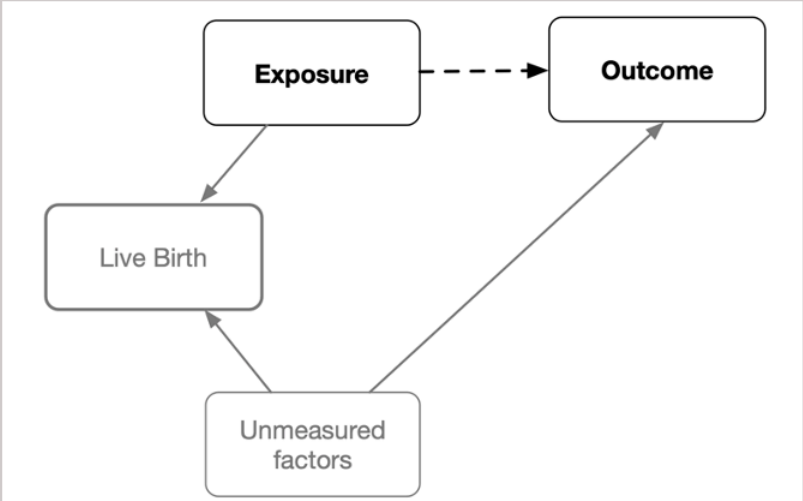

Figure 2.5 Directed acyclic graph illustrating potential live birth bias

A directed acyclic graph representing the bias resulting from conditioning on live birth, which leads to selection bias. Conditioning on the collider, *Live Birth,* opens the backdoor path:
Exposure → Live Birth ← Unobserved factors → Outcome

Adapted from: Raz R, Kioumourtzoglou MA, Weisskopf MG. Live-birth bias and observed associations between air pollution and autism. *Am J Epidemiol.* 2018; 187(11):2292-2296.

The "depletion of susceptibles" mechanism occurs in the presence of left truncation (also known as late or staggered entry); when the exposure preferentially results in the exclusion of fetuses (through a pregnancy loss) that are more susceptible to develop the outcome later on (268, 273). This essentially is the exposure preventing selection into the study sample (live-born infants that survive to the age of outcome diagnosis) of the susceptible fetuses because a live birth does not occur (214, 273, 276). Perinatal epidemiology studies of outcomes such as early pregnancy losses are vulnerable to left truncation when participants are already exposed prior to study initiation, such

that exposure-induced selection processes likely determine observational cohort formation from which exposure-outcomes effects are estimated (274). The "conditioning on a collider" mechanism occurs when the exposure and some other independent factors influence selection. Specifically, the exposure is associated with the outcome and preferentially results in pregnancy loss; and a second variable exists which is associated with both pregnancy loss and the outcome of interest. This second variable could be an unobserved, unknown, or simply ignored common cause of pregnancy loss and the outcome. By conditioning on pregnancy loss i.e. only including live born infants, a non-causal path is created between the exposure and the outcome via the second variable (268, 273).

The necessary data to mitigate live birth bias, namely data on pregnancy loss in the study population and/or all potential common causes between pregnancy loss and outcome of interest, are often not available (268). Consequently, unlike other types of selection bias such as the loss to follow-up or the "birthweight paradox", in the absence of this data this bias is less amenable to being addressed analytically because of the difficulty of adjustment for unobservable selection processes (274).

Immortal Time Bias

Exposure status follow-up times which are incorrectly handled in the analysis phase could introduce immortal time bias in favour of the exposed group (277, 278). This often takes place in studies in which the exposure is a treatment that is prescribed at variable times (with a delay or waiting period) after outcome diagnosis (279). Immortal time refers to period of follow-up in which, by design, the study outcome cannot occur in the exposed group because individuals ending up in the exposed group have to survive until the treatment definition is fulfilled (277). In pregnancy, this bias arises when entry into the exposed group depends on remaining pregnant long enough to have the opportunity for exposure or to start treatment (280). There are two scenarios in which immortal time bias occurs that can lead to either measurement bias or selection bias if immortal time is not appropriately accounted for in analysis. The first scenario is when the entire follow-up time of the exposed is assigned as exposed even though there were periods which were theoretically unexposed (before treatment was started) which can lead to misclassification (277, 278). The second scenario is when there is inappropriate exclusion of immortal times with the discarding of the unexposed follow-up periods of the exposed from the analysis which can lead to selection bias (196, 279, 281) (Figure 2.6). This can happen when the exposed cohort is formed from all participants who have received the treatment under study, thus they enter the cohort when treatment starts. All other

participants are considered unexposed with an arbitrarily defined cohort entry based on some 'entry event'. These two sets of participants are then followed up from cohort entry until the outcome event occurs. The exposed participants could also have had the same 'entry' event before their exposure-defined cohort entry time, so the period between this prior 'entry' event and the cohort entry for the exposed group is immortal and needs to be accounted for. 'Exposed' participants who previously had this entry event during the immortal period would have been classified in the unexposed group. Exclusion of this immortal person-time generated by the exposed participants introduces bias, because it should have been accounted for as unexposed. Ordinarily this period would add person-time to the denominator of the unexposed while not adding any new events to the numerator. By excluding this immortal time, the outcomes in the unexposed group will be artificially increased, because part of the denominator is missing so that the estimate is biased downwards (279, 282, 283).

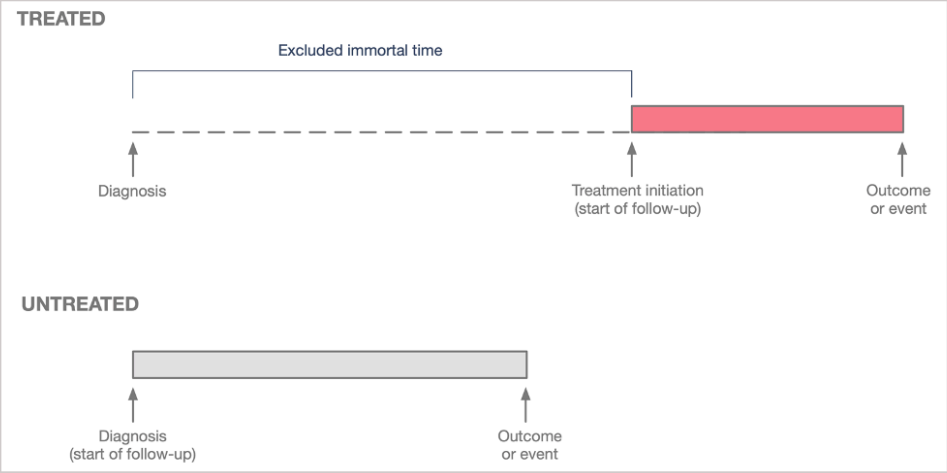

Figure 2.6 Illustration of immortal time bias

Exclusion of immortal time can lead to selection bias

Adapted from: Daniel S, Koren G, Lunenfeld E, Levy A. Immortal time bias in drug safety cohort studies: spontaneous abortion following nonsteroidal anti-inflammatory drug exposure. *Am J Obstet Gynecol.* 2015;212(3):307.e1-6.

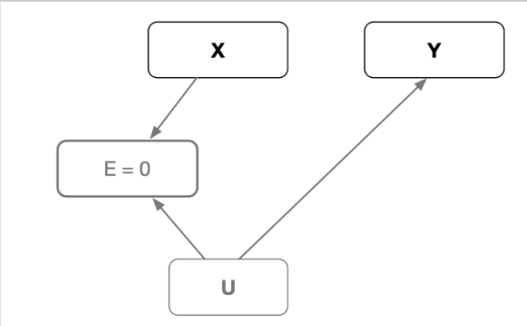

Figure 2.7 Directed acyclic graph illustrating selection bias due to immortal time exclusion

X = exposure, Y = outcome,
U = unmeasured protective cause of the outcome,
E = collider (exclusion of immortal times: E = 1 for excluded (immortal) person-times, E = 0 otherwise.)

Adapted from: Mansournia MA, Nazemipour M, Etminan M. Causal diagrams for immortal time bias. *Int J Epidemiol. 2021*; 50(5):1405-1409.

Insufficient consideration of immortal time can result in significant bias by conferring a spurious survival advantage to the treated group, which could potentially reverse the direction of the association (268, 277). The impact of this bias can be demonstrated in Figure 2.7. The DAG shows the null hypothesis of no association between the exposure and the outcome (no arrow from X→Y). With exclusion of the immortal time in analysis (by conditioning the collider), the exposure and outcome become associated because an open path has been created between the exposure and outcome (X→E←U→Y). Consequently, the

unexposed group (X = 0) with the unmeasured protective cause of the outcome (U = 1) is more likely to be excluded from the analysis (E = 1) than those without unmeasured protective cause (X = 0 and U = 0). This is because the exposed have a higher probability of the exposure than the latter, as they are more likely to survive the waiting time (279).

In an attempt to mitigate this selection bias, a person-time exposure definition that accounts for the immortal person-time as unexposed until the start of treatment can be used instead of the exclusion approach. This can be done by emulating the immortal times for the unexposed group from the distribution of exposed group immortal times, effectively recalibrating the follow-up times for the unexposed subjects. It should be noted that this can only partly remedy the selection bias problem in the exclusion scenario, so if the exposure is protective or has no effect the bias would be away from the null, and towards the null was a risk factor (279).

Conditioning on intermediates

In order to estimate direct effects of the exposure on the outcome without bias, the assumption of no unmeasured confounding for the exposure and outcome, and any intermediate variable and outcome needs to hold (195). Conditioning on an intermediate variable (I) (Figure 2.8a) or a descending proxy for an intermediate variable (Figure 2.8b), which can be both be viewed as a form of overadjustment, means that the total causal effect of X→Y association cannot be consistently estimated (255, 284). If the adjustment for the intermediate variable (or proxy/descendent variable) occurs in the presence of unmeasured confounder (U) between I→Y, then the X→Y relationship will be subject to "collider stratification bias" (Figure 2.8c) (197). Conditioning/adjustment can be done through restriction, stratification, matching (on the intermediary variable) or regression-based adjustment (285).

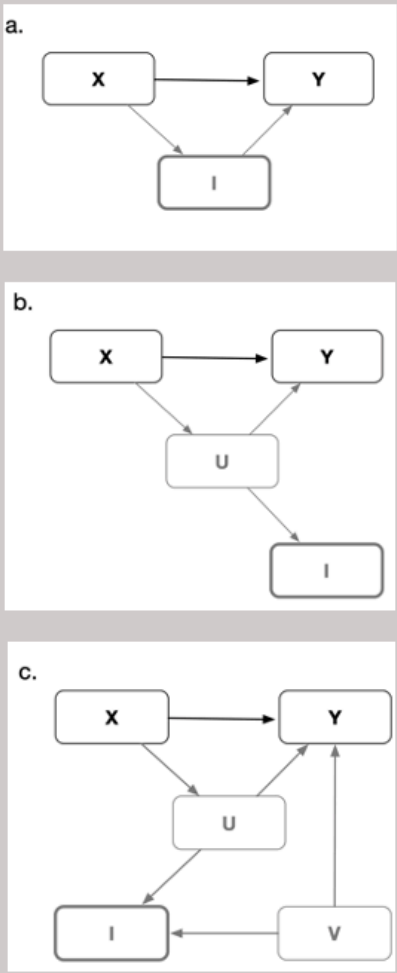

Figure 2.8 Scenarios for variables acting as intermediates

Adapted from: Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009; 20(4):488-495.

This is of particular concern in perinatal epidemiology, because of a unique tension that exists (286). On one hand there is the desire to obtain gestational age or birthweight specific associations (18, 287-289); while on the other there is an increased awareness that this conditioning action can result in severe biases (285). This issue was first highlighted in perinatal epidemiology research with the birthweight 'paradox' a controversial subject for many years (76).

In the relationship between maternal smoking and mortality in LBW infants, smoking appeared to be protective with LBW infants of smokers having lower mortality rates than those of non-smokers (18, 290-292). This was despite the widely accepted notion that infants of smokers had lower birthweights and were at increased neonatal mortality risk (76). Because of the accessibility of birthweight, data stratified on birthweight was frequently stratified on, which often produced a crossover of the birthweight-specific mortality curves (290). This action of conditioning on birthweight, a potential intermediate, induces collider bias that leads to an inverse association between the exposure and outcome among a subset of the infants, LBW infants in this case (268, 290). It is hypothesised that different pathways exist to achieve LBW, and only the LBW type caused by certain factors such as birth defects, and not others such as smoking increase mortality (76, 293) (Figure 2.9).

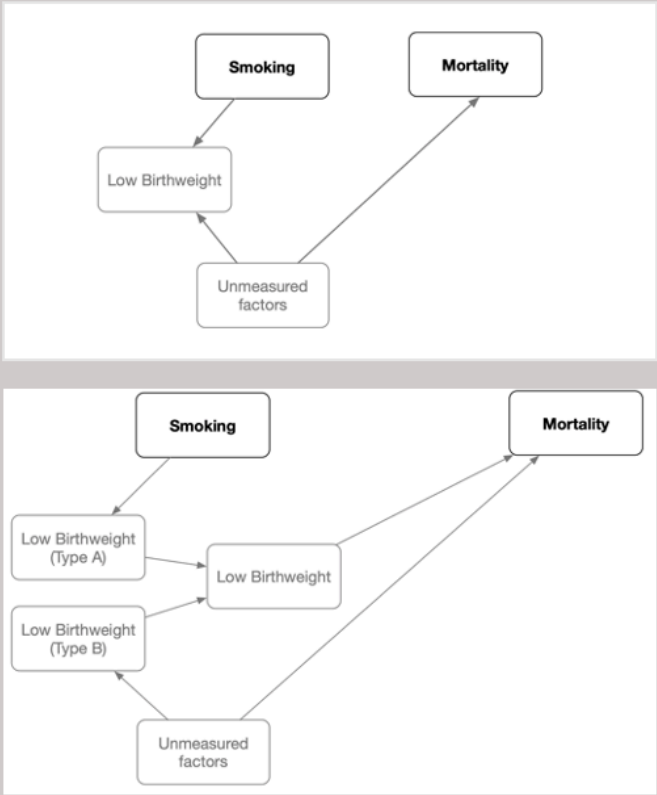

Figure 2.9 Causal structure of birthweight paradox

Adapted from: Hernández-Díaz S, Schisterman EF, Hernán MA. The birthweight "paradox" uncovered? *Am J Epidemiol.* 2006; 164(11):1115-1120.

This "paradox" is not limited to the smoking and infant mortality relationship. When the association between any pre-pregnancy variable and any postnatal outcome is being studied, conditioning on any variable affected by the exposure can introduce bias if the causal relation between the variables is similar to that represented in Figure 2.9 (290). While there is increased attention being paid to overadjustment and the potential for collider bias (particularly in relation to birthweight), many studies are still conditioning for gestational age, without clear rationale for doing so (294-297). This could be because gestational age is often conceptualised as a confounder because early delivery unquestionably contributes to morbidity and mortality (18). However, gestational age is a potential mediator of exposure–outcome relationship if timing of delivery is affected by the exposure in addition to the outcome (285). Conditioning will result in removal of some of the effect of interest or introduction of collider bias by opening the pathway between the exposure, intermediate, unmeasured factor, and the outcome (X→I←U→Y).

Preterm delivery as an endpoint is also appealing because in addition to being strongly associated with mortality, it is also a more common outcome than mortality. Despite this, it is typically conditioned on because preterm infants carry the burden of the factor/s that triggered early delivery (18); which coincidentally tend to be poorly characterised or unmeasured. Bias can be induced by the unmeasured factors that contribute to early delivery which also impact the outcome this makes PTD a collider. Conditioning on PTD blocks the effect of preterm on the outcome and opens a path between the exposure and outcome (Figure 2.10).

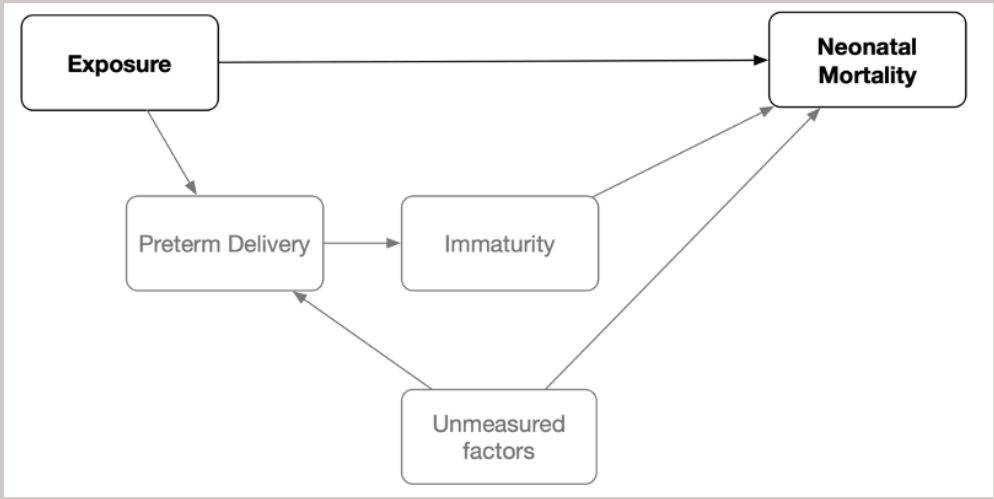

Figure 2.10 Causal association of preterm delivery and outcome

X" represents a measured variable of interest to the investigator
"U" represents unmeasured confounding variables

Adapted from: Wilcox AJ, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. *Am J Epidemiol.* 2011; 174(9):1062-1068.

Mitigation of this problem is through careful consideration of research question and the assumptions before decisions are taken about the necessity of conditioning on an intermediate (268). Conditioning on intermediates should not be done when the total effect or causal effects are the parameters of interest. Conversely, in cases where a specific pathway is of interest then conditioning could be warranted however this needs to be explicitly stated and the assumption that the effect of interest is more or less a direct effect requiring no unmeasured mediator–outcome confounding, needs to be met (268). If prediction rather than causation is of interest, restricting the population to specific subgroups of interest (e.g. preterm or low-birthweight categories) would not necessarily be problematic, and the estimated associations would still be of value (194, 268).

2.5 Discussion

2.5.1 Medication use in pregnancy

Pregnancy is a normal physiological process however it is often associated with pathological processes that necessitate the use of medications, which play important roles in promoting maternal well-being (194). There are however concerns around medication use in pregnancy because pregnancy physiology can affect the pharmacokinetics of medications used, and some medications can harm the fetus because of potential teratogenicity (298, 299). Despite these concerns, total avoidance of medication in pregnancy is not possible because some women enter pregnancy with co-morbid conditions that require ongoing and/or episodic treatment such as hypertension, diabetes mellitus, asthma, and epilepsy (298). Additionally, the presence of infections in pregnancy, for instance HIV, increases the need for anti-infective agents such as ART (194).

Given that events that occur during pregnancy have important influences on the outcome of pregnancy, as well as the health and wellbeing of newborns, it is essential to understand the safety of maternal medication use. However, safety of medications in pregnancy is often unknown/poorly understood, because of the ethical issues surrounding research in pregnancy linked to the uncertainties around the effects of medications on fetal development. Evidence on medication efficacy and safety, in the general population, is primarily based on randomised controlled trials. However, these trials routinely exclude pregnant women, with medication safety data coming from reproductive toxicity studies in animals and observational studies in humans (300). In the past, the majority of studies on the risks of maternal medication use were small field studies that had small sample sizes, lacked statistical power, or had sub-optimal study designs to pregnancy or infant outcomes (301). To circumvent these limitations, in recent years, this association has been assessed using data from larger cohort studies (302), and large national pregnancy medication exposure registries (303, 304) and administrative databases (305-307). While these new data sources have improved the data quality, this routinely collected data is not originally intended for research. So it is conceivable that bias, one of the many key challenges to making valid inferences, is becoming increasingly prevalent in the study of birth outcomes.

2.5.2 Example: Maternal antiretroviral use and adverse birth outcomes

The relationship between maternal ART use and adverse birth outcomes has been an area of research for many years, with conflicting results from multiple studies across both HICs and LMICs. Maternal HIV infection, in the pre-ART era, was associated with adverse birth outcomes including PTD, LBW and SGA infants and stillbirth; with the highest risk among women with advanced HIV disease or immunosuppression (308-313). The proposed mechanisms for these adverse birth

outcomes were poor maternal health, increased risk factors for co-infections and acute fetal HIV infection.

Provision of ART to pregnant and breastfeeding women has been successful in prevention of mother-to-child transmission of HIV (PMTCT) and improving maternal health and survival (314). However, despite these successes, concerns about increased risk of adverse birth outcomes among pregnant women living with HIV on ART were first raised in European studies (315, 316). Latin and North American studies, at the time, observed no associations(317-320). The majority of early data on this association came from cohorts in high income countries (HIC) (315, 321-324), with little to no data from LMICs where the majority of women requiring ART during pregnancy resided. The epidemiology of this association has been challenging, because of the competing forces that exist. ART may decrease some adverse birth outcomes by improving maternal health and reducing acute fetal infection. Conversely, ART has also been observed to increase the risk of these adverse outcomes, with risk greater than with nucleoside reverse transcriptase inhibitors (NRTI) monotherapy or dual therapy (325, 326). The proposed mechanisms for this increase, include ART-induced disruption or increased inflammatory response following ART initiation because of immune reconstitution syndrome (327).

Despite these ART-related increases of potential adverse outcomes, ART is unequivocally recommended in pregnancy because healthier women have better pregnancy and birth outcomes, and early initiation of ART is preferred for long-term health in any person living with HIV. While in the past HIV treatment research focused on vertical transmission and congenital anomalies, PMTCT programme successes and increasingly safe ART regimens have resulted in the focus shifting to other outcomes. These include PTD, which is the leading cause of neonatal and under-five mortality globally (328, 329); and has poor long-term outcomes in those that survive, particularly the very preterm infants (330, 331). Further, LBW and SGA infants are also at significant higher morbidity and mortality risk, particularly in LMIC (332, 333). High-quality data on the safety and efficacy of most antiretrovirals in pregnancy is limited, which is concerning because the presence and magnitude of the association between maternal ART use and adverse birth outcomes has been shown to differ according not just setting but also ART regimens and timing of initiation (334). In the PROMISE trial, triple drug ART was associated with an increased rate of preterm and low birthweight infants compared to zidovudine (ZDV) monotherapy; with those randomised to ritonavir-boosted lopinavir (LPV/r) + tenofovir disoproxil fumarate (TDF) having a significantly higher risk of adverse outcomes (325). These findings highlighted the need to focus on the role of ART regimens, in particular protease inhibitors (PIs). Prior to PROMISE, PIs were implicated in the increased risk of adverse birth outcomes in both observational (335, 336) and randomised (337) studies. The

proposed mechanism of action is linked to adrenal system interference implicated in the spontaneous onset of labour (323) and/or reductions in progesterone levels during pregnancy which could affect fetal growth (338). In more recent times, the Tsepamo study in Botswana, also showed that birth outcomes differed by maternal ART regimen (336). Women on TDF + emtricitabine (FTC) + efavirenz (EFV) had the lowest proportion of adverse outcomes compared to those on other ART regimens which included ZDV + lamivudine (3TC) + nevirapine (NVP) and ZDV+3TC+LPV/r (336, 339). This study based on national surveillance data found that severe outcomes such as stillbirth also differed by regimen, with an increased risk among women on ZDV+3TC+NVP compared with other ART regimens (336).

While the older more toxic regimens are no longer in use, early experiences with these regimens need to be considered when dealing with contemporary maternal ART regimens. In 2018, WHO treatment guidelines recommend the replacement of EFV in the first-line treatment with dolutegravir (DTG), an integrase inhibitor that is potent, well tolerated and has a high genetic barrier to resistance (340). The VESTED trial investigated the safety and efficacy of newer ART regimens, likely to be used by many women initiating ART during pregnancy, but for which limited rigorous pregnancy safety and efficacy data exist (341). The DTG + FTC + tenofovir alafenamide fumarate (TAF) arm of the trial had the lowest proportion of composite adverse birth outcomes compared to the other regimens (DTG + FTC +TDF and TDF + FTC + EFV); primarily driven by lower rates of preterm and to a lesser extend small for SGA infants in this group (341). This collectively underscores the importance of evaluating birth outcomes with newer ART regimens.

Timing of ART initiation with respect to gestation also appears to be critical, however there is limited data, so uncertainties remain around its role in the occurrence of increased risks of adverse birth outcomes (334). With increasing numbers of women conceiving on ART, more information is required on the safety of regimens and the impact of timing of initiation. A meta-analysis showed that women on preconception ART were at increased risk of delivering preterm, very preterm and LBW infants, compared to those who initiated during pregnancy (342). The data on the role of timing of ART initiation are not conclusive, with some studies finding no difference (343, 344). These results may depend on regimens used, as demonstrated by an analysis of subsequent pregnancies in PROMISE. Women on their second pregnancies who were now conceiving on ART (mostly PI) had a higher proportion of miscarriages and stillbirth compared to women who initiated during pregnancy (345). There are concerns previous studies investigating the timing of ART initiation and adverse birth outcomes did not account for gestational age at ART initiation, which could induce selection bias because of the unequal opportunities of experiencing the outcome in women who initiating ART later in pregnancy compared to those who initiated earlier or before pregnancy (346). Based on

these selection concerns a simulation trial that assessed the impact of the exclusion of women delivering pre-ART initiation, showed that the exclusion decreased the risk of women included in studies investigating this association (347). When the simulated randomised data was compared to an analysis that mimicked previous observational studies, the randomised data showed no increased risk with preconception initiation, while the other analysis showed an increased risk with preconception initiation (347).

Interpretation of observational data on ART safety is complicated by changing treatment guidelines and comparator groups, low quality gestational age data in some studies and other bias related concerns that are common with all observational research. In the treat-all era, as more women will conceive on ART, high quality data are required that are relevant to the changing population of pregnant women living with HIV and to new antiretroviral drugs, and that address limitations in the existing evidence base (334).

2.5.2.1 Bias related concerns

Putative explanations for the conflicting and inconsistent findings across different settings, in the association between maternal ART use and adverse birth outcomes, have been linked to numerous factors linked to bias and confounding. These include bias due to differences in study designs (case ascertainment and inclusion criteria), study populations (different baseline rates of outcomes of interest) and analytical approaches (differences in choices for reference groups, lack of adjustment for risk factors) (Figure 2.11).

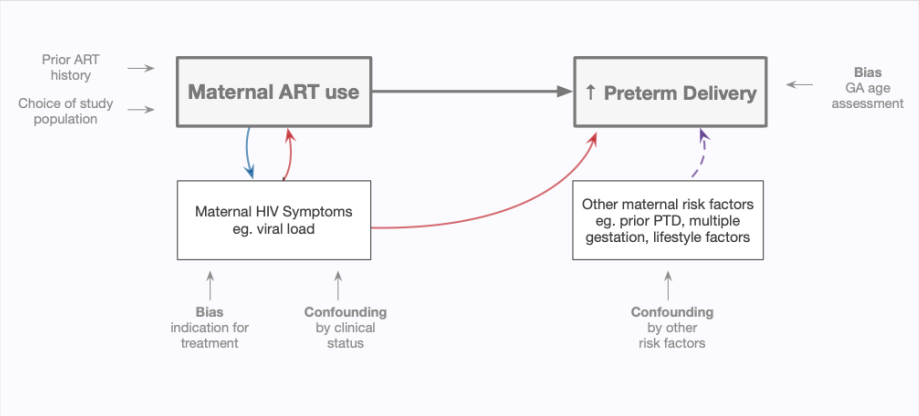

Figure 2.11 Bias related concerns in the association between maternal ART use and adverse birth outcomes

Other bias-related concerns such as differences in the quality of measurement of exposure timing and outcome ascertainment persist in perinatal epidemiology generally and specifically when investigating this association. This is in part because in the majority of studies investigating this association, the birth outcomes analyses tend to be planned or unplanned secondary analyses.

Consequently, data collection procedures are not designed to examine varies aspects of this association and so the best high-quality measures and procedures are not always utilised. This predominantly the case with gestational age assessment, where measurement error due to menstrual dating and/or clinical examination may introduce bias where an imprecise gestational age can result in outcome misclassification. Given that many adverse birth outcomes are defined by gestational duration, it is critical that high quality measures of gestation are used to ensure an accurate degree of exposure and outcome ascertainment.

The definition of exposure and the presence of exposure misclassification is linked to research objectives, if the objective is to evaluate association of any ART use in pregnancy, then the exposure assessment based on self-report may not suffer from much misclassification - measurement error could still occur however estimates would be unbiased (under the null hypothesis). In contrast, when the objective is to evaluate more nuanced aspects of the association such as duration of ART use or specific regimens, misclassification with biased estimates is likely if there is a reliance on self-report. Even when exposure information is available from other sources such as medical records or biological assays, unless data is collected specifically for the purposes of the study there are no assurances that the data is any more accurate than self-report (5). Additionally, this association is also affected when information on participant characteristics is missing or unavailable, as is common in observational data. This can be seen as an extreme form of measurement error, which impacts the ability to adjust for important confounders such as demographics, substance use, HIV factors (degree of immunosuppression, indication for treatment) and coinfection. When assessing this association, there are several factors linked to the exposure and outcomes and each of these can potentially be measured incorrectly and thus affect the findings. The exposure maternal ART use is a function of the regimen type used, duration of use and adherence (Figure 2.12); and each of these factors independently contribute to the risk of an outcome of interest. Misclassification of these factors can occur leading to incorrect assignment to an exposure group, which could lead to incorrect inferences. For instance, in the case

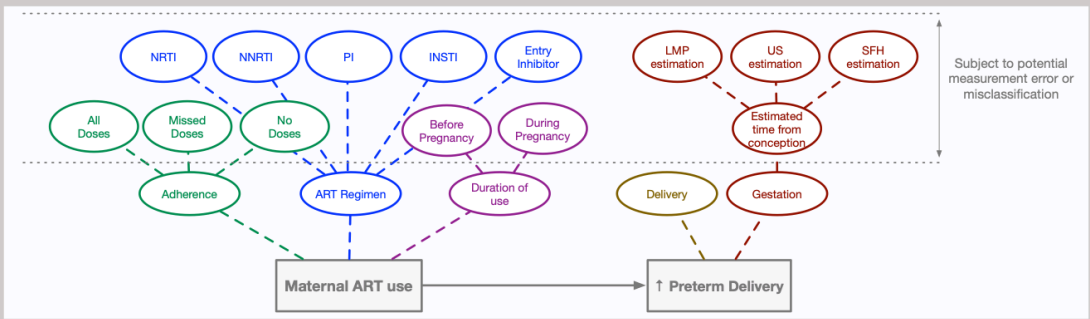

Figure 2.12 Potential sources of measurement error in association between maternal ART use and preterm delivery

NRTI - Nucleoside/nucleotide reverse transcriptase inhibitors, NNRTI - Non-nucleoside reverse transcriptase inhibitors, PI – protease inhibitor, INSTI - Integrase strand transfer inhibitors, LMP – last menstrual period, US – ultrasound, SFH – symphysis fundal height

of a women who is classified as being on ART, based on self-report or because clinic records indicate ART use, is not adherent and has missed all her doses – if she were to deliver preterm this would be unrelated to ART however this could incorrectly be attributed to ART use. Preterm delivery as an outcome is a function of a delivery event and the gestation at which this delivery occurred. The delivery event is not subject to error, however the GA at which this delivery event occurred which determines classification of the delivery as preterm or not is subject to measurement error and the type and magnitude depends on the GA assessment method used.

2.5.3 Need for future research

From a public health perspective, the ability to generalise causal effects from both observational and randomised studies is fundamental. To achieve this, high-quality data are essential so that evidence from well-designed studies can be incorporated in public health decision making and treatment guidelines. This will require optimisation of current study designs and analytical techniques to study adverse birth outcomes and their risk factors more accurately, by taking into account the important methodological considerations briefly described in this review. While well-designed randomised trials remain the scientific gold standard, they are frequently limited by sample sizes lacking the statistical power to detect changes in important maternal, fetal, or neonatal outcomes. They typically do not target the enrolment of "conception" cohort, with the trials tending to randomise after the first trimester (>14 weeks). Additionally, it is not always feasible to incorporate randomisation or blinding into research involving pregnant women (348). Consequently, birth outcome is typically investigated through the analysis of existing data from observational studies or routinely collected data. While this may produce information that accurately reflects maternity care processes and outcomes, this data is not specifically focused on birth outcomes. This likely affects data quality, particularly related to measurement error in gestational age which is an important variable for all outcomes discussed in this review.

In the context of aetiological research, the common view is that random measurement error will result in attenuation (bias towards the null) of the estimated exposure-outcome association. A recent systematic review found that 'attenuation of effect' remains a prevailing notion and this almost always remains unsubstantiated by statistical analyses (244). Under this assumption of effect attenuation, estimated exposure-outcome associations are considered conservative estimates of the counterfactual scenario where measurement error is absent. Paradoxically this leads to the notion that any significant findings achieved in the presence of measurement error provide evidence that the underlying effect is strong and robust (349), even though statistical significance conveys little information when there is measurement error since poor measurement can exaggerate effect size

estimates. It has now been demonstrated that random measurement error in the exposure and/or confounder does not automatically result in attenuation, with situations where exaggeration has occurred (350). This highlights the difficulty in estimating the magnitude and direction of measurement error related bias without applying specialised analytical methods. Various methods have been proposed that provide quantitative estimates of the direction, magnitude and uncertainty arising from systematic errors; as well as enabling identification of sources of systematic error and their quantification (260). With the increasing use of routinely collected data, renewed attention for the complex impact of measurement error and approaches for dealing with measurement error in perinatal epidemiology are vital. In particular, how to account for measurement error during analysis, when it cannot be prevented during data collection. Despite its importance, it is unclear to what extent measurement error related issues are currently considered in research. Measurement error correction methods have been developed, which could be used to address gestational age measurement error (351). It is however evident from the literature that these correction methods are currently performed to a very limited degree particularly in relation to birth outcomes research. The impact of endogenous selection bias has not appeared to be a concern of equal importance to that of confounding and measurement error/misclassification. This is likely reflection of the difficulty in investigating the impact of selection bias on internal validity compared to that of other biases; particularly amidst the dynamic, hidden, and interconnected populations of pregnancy, birth, and childhood (266). Despite the potential for selection bias studies, it is not uncommon for results to be considered as though internally valid and generalised to a (usually undefined) target population. To achieve generalisability, the first step is to identify causal effects in the presence of selection bias, followed by the application of appropriated analytical methods (264). However, while there has been a movement towards quantifying potential bias in epidemiology studies, this is difficult to accomplish for selection bias because, by definition, data on those not included are unavailable (260). Further, there are few analyses of associations between survival to livebirth and potential exposures to inform quantitative bias analyses (260, 264, 352, 353). So while adjusting away selection may not always be possible in perinatal epidemiology, vigilance will be required in seeking to understand these processes and how they affect our results. Analytical approaches need to be formulated that conform to the foundational tenets of epidemiology. This involves confronting bias related to selection processes and competing risks, in order to understand and address these biases (266).

A better understanding and quantification of the error introduced when using commonly used, but error prone, assessment methods, is necessary so that adverse birth outcome rates can be appropriately adjusted for when necessary. Additionally, a nuanced understanding of data limitations is required, without which, results of global estimates or aetiologic research need to be

interpreted with caution to avoid incorrect conclusions. New data analysis approaches should be explored as they may provide more efficient ways to use existing data to inform policy and practice.

2.5.4 Conclusion

In conclusion, this review has summarised the important knowledge about the epidemiological measures of pregnancy and the various interlinked outcomes that span from conception to birth, specifically pregnancy loss, gestational duration, and fetal growth. In general, the importance of bias in epidemiology is well known; however more high-quality evidence is needed about methodological concerns, related to measurement and selection, that may arise during the study of pregnancy and birth outcomes and potential mitigation strategies.

Using quantitative data from a combination of well-designed cohort studies, a pregnancy exposure registry and provincial health database, this thesis aims to contribute to the evidence base related to the association between maternal ART use and adverse birth outcomes. This will be done while addressing some of the limitations inherent in research of this association to date. Specifically, gestational age measurement error will be assessed by assessment method, to determine the extent and impact of bias from the misclassification of preterm versus term versus post-term deliveries. Building on this the aim is to explore imputation based and estimation based bias correction methods to reduce the impact of this bias on findings of this association. The impact of study population selection from an underlying population alongside differing quality of variables of interest on the association will also be explored. Lastly, in order to better inform understanding of underlying mechanisms, this thesis intends to produce evidence that can be used to inform future research methodology, through the examination of alternative analytical approaches for improving the accuracy of exposure and/or risk factor status classification to avoid misclassification which could introduce bias and affect study inferences.

2.6 References

1. Wilcox AJ. Fertility and Pregnancy: An Epidemiologic Persective. Oxford University Press: New York, USA; 2010.
2. Cha J, Sun X, Dey SK. Mechanisms of implantation: strategies for successful pregnancy. *Nat Med.* 2012;18(12):1754-1767.
3. Norwitz ER, Schust DJ, Fisher SJ. Implantation and the survival of early pregnancy. *N Engl J Med.* 2001;345(19):1400-1408.
4. Silver RM, Branch DW, Goldenberg R, Iams JD, Klebanoff MA. Nomenclature for pregnancy outcomes: time for a change. *Obstet Gynecol*. 2011;118(6):1402-1408.
5. Slama R, Ballester F, Casas M, et al. Epidemiologic tools to study the influence of environmental factors on fecundity and pregnancy-related outcomes. *Epidemiol Rev.* 2014;36(1):148-164.
6. Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic measures of the course and outcome of pregnancy. *Epidemiol Rev*. 2002;24(2):91-101.
7. Cole T. Growth and organ development. Breast-Feeding: Early Influences on Later Health. 2009:1-13.
8. Richardson A, Gallos I, Dobson S, et al. Accuracy of first-trimester ultrasound in diagnosis of intrauterine pregnancy prior to visualization of the yolk sac: a systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2015;46(2):142-149.
9. Falkner F, Tanner JM. Human Growth: A Comprehensive Treatise Volume 1 Developmental Biology Prenatal Growth. Springer: New York, USA; 2012.
10. Korones S. Anatomic aspects of fetal development. *Glob Libr Womens Med.* 2008;1:1-22.
11. Butt K, Lim K, Bly S, et al. Determination of gestational age by ultrasound. *J Obstet Gynaecol Can*. 2014;36(2):171-181.
12. MacGregor SN, Sabbagha R. Assessment of gestational age by ultrasound. *Glob Libr Womens Med.* 2008;1:1-18.
13. Hadlock FP, Harrist RB, Shah YP, et al. Estimating fetal age using multiple parameters: a prospective evaluation in a racially mixed population. *Am J Obstet Gynecol*. 1987;156(4):955-957.
14. Hadlock FP, Deter RL, Harrist RB, Park S. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. *Radiology*. 1984;152(2):497-501.
15. Lockwood CJ, Weiner S. Assessment of fetal growth. *Clin Perinatol.* 1986;13(1):3-35.
16. Karaca NH. Prenatal Development. Handbook of Research on Prenatal, Postnatal, and Early Childhood Development. pg 140-58. IGI Global: Pennsylvania, USA; 2020.
17. Buck Louis GMP, Platt RW. Reproductive and perinatal epidemiology. Oxford University Press: New York, USA; 2011.

18. Wilcox AJ, Weinberg CR, Basso OJ. On the pitfalls of adjusting for gestational age at birth. *Am J Epidemiol.* 2011;174(9):1062-1068.
19. Smith GC. Quantifying the risk of different types of perinatal death in relation to gestational age: researchers at risk of causing confusion. *Paediatr Perinat Epidemiol.* 2016;30(1):18-19
20. Blair E, Liu Y, Cosgrove P. Choosing the best estimate of gestational age from routinely collected population-based perinatal data. *Paediatr Perinat Epidemiol.* 2004;18(4):270-276.
21. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet.* 2008;371(9608):261-269.
22. Cotzias CS, Paterson-Brown S, Fisk NM. Prospective risk of unexplained stillbirth in singleton pregnancies at term: population based analysis. *BMJ*. 1999;319(7205):287-288.
23. Caughey AB, Snegovskikh VV, Norwitz ER. Postterm pregnancy: how can we improve outcomes? *Obstet Gynecol Surv.* 2008;63(11):715-724.
24. Baird DD, Weinberg CR, Wilcox AJ, McConnaughey DR, Musey PI. Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation. *Stat Med.* 1991;10(2):255-266.
25. Kesner JS, Knecht EA, Krieg EF Jr, Wilcox AJ, O'Connor JF. Detecting pre-ovulatory luteinizing hormone surges in urine. *Hum Reprod.* 1998;13(1):15-21.
26. Jukic AM, Baird DD, Weinberg CR, McConnaughey DR, Wilcox AJ. Length of human pregnancy and contributors to its natural variation. *Hum Reprod.* 2013;28(10):2848-2855.
27. Wilcox AJ, Dunson D, Baird DD. The timing of the “fertile window” in the menstrual cycle: day specific estimates from a prospective study. *BMJ*. 2000;321(7271):1259-1262.
28. Blondel B, Morin I, Platt RW, et al. Algorithms for combining menstrual and ultrasound estimates of gestational age: consequences for rates of preterm and postterm birth. *Br J Obstet Gynaecol*. 2002;109(6):718-720.
29. WHO. WHO recommendation on symphysis-fundal height measurement. Geneva, Switzerland: World Health Organisation; 2016.
30. Vik T, Vatten L, Markestad T, Jacobsen G, Bakketeig LS. Dubowitz assessment of gestational age and agreement with prenatal methods. *Am J Perinatol.* 1997;14(6):369-373.
31. Verhoeff FH, Milligan P, Brabin BJ, Mlanga S, Nakoma V. Gestational age assessment by nurses in a developing country using the Ballard method, external criteria only. *Ann Trop Paediatr.* 1997;17(4):333-342.
32. Dubowitz LM, Dubowitz V, Goldberg C. Clinical assessment of gestational age in the newborn infant. *J Pediatr.* 1970;77(1):1-10.
33. Ballard JL, Khoury JC, Wedig K, et al. New Ballard Score, expanded to include extremely premature infants. *J Pediatr.* 1991;119(3):417-423.

34. Clark SL, Perlin JB, Fraker S, et al. Association of obstetric intervention with temporal patterns of childbirth. *Obstet Gynecol.* 124(5):873-880.
35. Savitz DA, Blackmore CA, Thorp JM. Epidemiologic characteristics of preterm delivery: etiologic heterogeneity. *Am J Obstet Gynecol*. 1991;164(2):467-471.
36. Kramer MS, Demissie K, Yang H, et al. The contribution of mild and moderate preterm birth to infant mortality. 2000;284(7):843-849.
37. Kramer MS, Papageorghiou A, Culhane J, et al. Challenges in defining and classifying the preterm birth syndrome. *Am J Obstet Gynecol.* 2012;206(2):108-112.
38. Stout MJ, Busam R, Macones GA, Tuuli MG. Spontaneous and indicated preterm birth subtypes: interobserver agreement and accuracy of classification. *Am J Obstet Gynecol*. 2014;211(5):530.e1-e4.
39. Ananth CV, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J Matern Fetal Neonatal Med*. 2006;19(12):773-782.
40. Manuck TA, Esplin MS, Biggio J, et al. The phenotype of spontaneous preterm birth: application of a clinical phenotyping tool. *Am J Obstet Gynecol*. 2015;212(4):487.e1-e11.
41. WHO. Recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. Geneva, Swizterland: World Health Organization; 1977.
42. Shea KM, Wilcox AJ, Little RE. Postterm delivery: a challenge for epidemiologic research. *Epidemiology.* 1998;9(2):199-204.
43. Baird DD, McConnaughey DR, Weinberg CR, et al. Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. *Epidemiology.* 1995;6(5):547-550.
44. Galal M, Symonds I, Murray H, Petraglia F, Smith R. Postterm pregnancy. *Facts Views Vis Obgyn.* 2012;4(3):175-187.
45. Laursen M, Bille C, Olesen AW, et al. Genetic influence on prolonged gestation: a population-based Danish twin study. *Am J Obstet Gynecol*. 2004;190(2):489-494.
46. Doherty L, Norwitz ER. Prolonged pregnancy: when should we intervene? *Curr Opin Obstet Gynecol*. 2008;20(6):519-527.
47. Mogren I, Stenlund H, Högberg U. Recurrence of prolonged pregnancy. *Int J Epidemiol.* 1999;28(2):253-257.
48. Tang AW, Agarwal U. Fetal growth disorders. In Kumar B, Alfirevic Z (Eds). Fetal Medicine (Royal College of Obstetricians and Gynaecologists Advanced Skills. pg V-Vi). Cambridge University Press: Cambridge, UK; 2016

49. Singer DB, Wigglesworth JS. Fetal growth and maturation: with standards for body and organ development. In: Singer DB, Wigglesworth JS (Eds). Textbook of Fetal and Perinatal Pathology. 2nd ed. pg 8-40. Blackwell Science: Oxford, UK; 1998.
50. Hay Jr WW. Intrauterine growth restriction. Perinatal Nutrition. pg 123-164. CRC Press: Florida, USA; 2004.
51. Mifsud W, Sebire NJ. Placental pathology in early-onset and late-onset fetal growth restriction. *Fetal Diagn Ther*. 2014;36(2):117-128.
52. Kingdom J, Huppertz B, Seaward G, Kaufmann P. Development of the placental villous tree and its consequences for fetal growth. *Eur J Obstet Gynecol Reprod Biol.* 2000;92(1):35-43.
53. Damhuis SE, Ganzevoort W, Gordijn SJ. Abnormal Fetal Growth: Small for Gestational Age, Fetal Growth Restriction, Large for Gestational Age: Definitions and Epidemiology. *Obstet Gynecol Clin North Am*. 2021;48(2):267-279.
54. Aviram A, Sherman C, Kingdom J, et al. Defining early vs late fetal growth restriction by placental pathology. *Acta Obstet Gynecol Scand.* 2019;98(3):365-373.
55. Peleg D, Kennedy CM, Hunter SK. Intrauterine growth restriction: identification and management. *Am Fam Physician.* 1998;58(2):453-460.
56. Cox P, Marton T. Pathological assessment of intrauterine growth restriction. *Best Pract Res Clin Obstet Gynaecol*. 2009;23(6):751-764.
57. Guerby P, Bujold E. Early detection and prevention of intrauterine growth restriction and its consequences. *JAMA Pediatr.* 2020;174(8):749-750.
58. Murmu S, Jaiswal A. A Study to Evaluate the Clinical Profiles of Asymmetric and Mixed Types of Intrauterine Growth Retardation. *Int J Sci Study*. 2019;7(9):5-9.
59. Walker D, Marlow N. Neurocognitive outcome following fetal growth restriction. *Arch Dis Child Fetal Neonatal Ed*. 2008;93(4):F322-F5.
60. Flenady V, Koopmans L, Middleton P, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *Lancet*. 2011;377(9774):1331-1340.
61. Burton GJ, Fowden AL, Thornburg KL. Placental origins of chronic disease. *Physiol Rev*. 2016;96(4):1509-1565.
62. Van Assche A. Symmetric and asymmetric fetal macrosomia in relation to long-term consequences. *Am J Obstet Gynecol*. 1997;177(6):1563-1564.
63. Fowler GC. Obstetric Ultrasound. Fowler GC (Ed). Elsevier Health Sciences: New York, USA; 2019.
64. Schwartz R. Hyperinsulinemia and macrosomia. *Mass Medical Soc*. 1990;323(5):340-342.
65. Farquhar JW. The Pregnant Diabetic and her Newborn: Problems and Management. *Pediatrics*. 1968;42(2):383-389.

66. Henriksen T. The macrosomic fetus: a challenge in current obstetrics. *Acta Obstet Gynecol Scand*. 2008;87(2):134-145.
67. Ju H, Chadha Y, Donovan T, O'Rourke P. Fetal macrosomia and pregnancy outcomes. *Aust N Z J Obstet Gynaecol*. 2009;49(5):504-509.
68. Carter EB, Stockburger J, Tuuli MG, et al. Large-for-gestational age and stillbirth: is there a role for antenatal testing? *Ultrasound Obstet Gynecol*. 2019;54(3):334-347.
69. Stock SJ, Myers J. Defining abnormal fetal growth and perinatal risk: population or customized standards? *PLoS Med.* 2017;14(1):e1002229.
70. Ohuma EO, Villar J, Feng Y, et al. Fetal growth velocity standards from the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Am J Obstet Gynecol*. 2021;224(2):208.e1-e18.
71. Altman DG, Hytten FE. Intrauterine growth retardation: let's be clear about it. *Br J Obstet Gynaecol.* 1989;96(10):1127-1128.
72. Salomon L. Early fetal growth: concepts and pitfalls. pg 385-389. Wiley Online Library: New York, USA; 2010.
73. Anderson N, Jolley I, Wells J. Sonographic estimation of fetal weight: comparison of bias, precision and consistency using 12 different formulae. *Ultrasound Obstet Gynecol*. 2007;30(2):173-179.
74. Arduini D, Giacomello F. Fetal biometry, estimation of gestational age, assessment of fetal growth. In: Wladimiroff JW, Eik-Nes SH (Eds). *Ultrasound Obstet Gynecol*. Elsevier: New York, USA; 2009
75. Stirnemann J, Salomon LJ, Papageorghiou A. The reality of estimating fetal weight (EFW): expectations and weaknesses. *Ultrasound Obstet Gynecol*. 2019.
76. Wilcox AJ. On the importance—and the unimportance—of birthweight. *Int J Epidemiol.* 2001;30(6):1233-1241.
77. Cutland CL, Lackritz EM, Mallett-Moore T, et al. Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. *Vaccine.* 2017;35(48 Part A):6492-6500.
78. Umbach DM, Wilcox AJ. A technique for measuring epidemiologically useful features of birthweight distributions. *Stat Med*. 1996;15(13):1333-1348.
79. Hemachandra AH, Klebanoff MA. Use of serial ultrasound to identify periods of fetal growth restriction in relation to neonatal anthropometry. *Am J Hum Biol*. 2006;18(6):791-797.
80. Hutcheon JA, Liauw J. Should fetal growth charts be references or standards? *Epidemiology.* 2021;32(1):14-17.
81. Mayer C, Joseph K. Fetal growth: a review of terms, concepts and issues relevant to obstetrics. *Ultrasound Obstet Gynecol*. 2013;41(2):136-145.

82. Kiserud T, Piaggio G, Carroli G, et al. The World Health Organization fetal growth charts: a multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. *PLoS Med*. 2017;14(1):e1002220.

83. Papageorghiou AT, Ohuma EO, Altman DG, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet.* 2014;384(9946):869-879.

84. WHO. International statistical classification of diseases and related health problems. Geneva, Switzerland: World Health Organization; 2004.

85. UNICEF. The state of the world's children 2008: Child survival. New York, US: The United Nations Children's Fund (UNICEF); 2007.

86. Hughes MM, Black RE, Katz J. 2500-g low birth weight cutoff: history and implications for future research and policy. *Matern Child Health J.* 2017;21(2):283-289.

87. Sacks DA. Etiology, detection, and management of fetal macrosomia in pregnancies complicated by diabetes mellitus. *Clin Obstet Gynecol*. 2007;50(4):980-989.

88. Ye J, Torloni MR, Ota E, et al. Searching for the definition of macrosomia through an outcome-based approach in low-and middle-income countries: a secondary analysis of the WHO Global Survey in Africa, Asia and Latin America*. BMC Pregnancy Childbirth*. 2015;15(1):1-10.

89. Boyd ME, Usher RH, Mclean FH. Fetal macrosomia: prediction, risks, proposed management. *Obstet Gynecol.* 1983;61(6):715-722.

90. Zhang X, Decker A, Platt RW, Kramer MS. How big is too big? The perinatal consequences of fetal macrosomia. *Am J Obstet Gynecol*. 2008;198(5):517.e1-e6.

91. Lim JH, Tan BC, Jammal AE, Symonds E. Delivery of macrosomic babies: management and outcomes of 330 cases. *J Obstet Gynaecol.* 2002;22(4):370-374.

92. Spellacy W, Miller S, Winegar A, Peterson P. Macrosomia--maternal characteristics and infant complications. *Obstet Gynecol.* 1985;66(2):158-161.

93. Stotland N, Caughey A, Breed E, Escobar G. Risk factors and obstetric complications associated with macrosomia. *Int J Gynaecol Obstet*. 2004;87(3):220-226.

94. Haram K, Pirhonen J, Bergsjø P. Suspected big baby: a difficult clinical problem in obstetrics. *Acta Obstet Gynecol Scand*. 2002;81(3):185-94.

95. Martorell R, Khan LK, Hughes ML, Grummer-Strawn LM. Obesity in women from developing countries. *Eur J Clin Nutr*. 2000;54(3):247-252.

96. Roglic G. Diabetes in women: the global perspective. *Int J Gynaecol Obstet*. 2009;104(Suppl 1):S11-S3.

97. Zhang Y, Li H, Liu SJ, et al. The associations of high birth weight with blood pressure and hypertension in later life: a systematic review and meta-analysis. *Hypertens Res*. 2013;36(8):725-735.

98. Savona-Ventura C, Chircop M. Birth weight influence on the subsequent development of gestational diabetes mellitus. *Acta Diabetol.* 2003;40(2):101-104.

99. Lee PA, Chernausek SD, Hokken-Koelega AC, Czernichow P. International Small for Gestational Age Advisory Board consensus development conference statement: management of short children born small for gestational age, April 24–October 1, 2001. *Pediatrics.* 2003;111(6 Pt 1):1253-1261.

100. Battaglia FC, Lubchenco LO. A practical classification of newborn infants by weight and gestational age. *J Pediatr.* 1967;71(2):159-163.

101. Clayton P, Cianfarani S, Czernichow P, et al. Management of the child born small for gestational age through to adulthood: a consensus statement of the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society. *J Clin Endocrinol Metab*. 2007;92(3):804-810.

102. Lee AC, Katz J, Blencowe H, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Glob Health*. 2013;1(1):e26–36.

103. Schlaudecker EP, Munoz FM, Bardají A, et al. Small for gestational age: Case definition & guidelines for data collection, analysis, and presentation of maternal immunisation safety data. *Vaccine.* 2017;35(48 Part A):6518-6528.

104. Osuchukwu OO, Reed DJ. Small for Gestational Age. In: StatPearls [Internet]. Florida, USA: StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK563247/ (accessed 15 January 2022)

105. Villar J, de Onis M, Kestler E, et al. The differential neonatal morbidity of the intrauterine growth retardation syndrome. *Am J Obstet Gynecol.* 1990;163(1):151-157.

106. Ido P, Gale R, Laor A, et al. The cognitive outcome of full-term small for gestational age infants at late adolescence. *Obstet Gynecol*. 1995;85(3):452-456.

107. Saenger P, Czernichow P, Hughes I, Reiter EO. Small for gestational age: short stature and beyond. *Endocr Rev*. 2007;28(2):219-251.

108. Karlberg J, Albertsson-Wikland K. Growth in full-term small-for-gestational-age infants: from birth to final height. *Pediatr Res.* 1995;38(5):733-739.

109. Westwood M, Kramer MS, Munz D, Lovett JM, Watters GV. Growth and development of full-term nonasphyxiated small-for-gestational-age newborns: follow-up through adolescence. *Pediatrics*. 1983;71(3):376-382.

110. Larroque B, Bertrais S, Czernichow P, Léger J. School difficulties in 20-year-olds who were born small for gestational age at term in a regional cohort study. *Pediatrics.* 2001;108(1):111-115.

111. Leger J, Levy-Marchal C, Bloch J, et al. Reduced final height and indications for insulin resistance in 20 year olds born small for gestational age: regional cohort study. *BMJ*. 1997;315(7104):341-347.
112. Carberry AE, Gordon A, Bond DM, et al. Customised versus population-based growth charts as a screening tool for detecting small for gestational age infants in low-risk pregnant women. *Cochrane Database Syst Rev.* 2014;2014(5):CD008549
113. Silasi M. Fetal Macrosomia. Obstetric Imaging: Fetal Diagnosis and Care. pg 460-462. Elsevier: New York, USA; 2018.
114. Ye J, Zhang L, Chen Y, et al. Searching for the definition of macrosomia through an outcome-based approach. *PLoS One*. 2014;9(6):e100192.
115. Xu H, Simonet F, Luo ZC. Optimal birth weight percentile cut-offs in defining small-or large-for-gestational-age. *Acta Paediatr.* 2010;99(4):550-555.
116. Simmons R. Abnormalities of fetal growth. Avery's diseases of the newborn. pg 61-69. Elsevier: New York, USA; 2018.
117. Bonnett T, Furara S, Allgar V, Roberts L. Management of large-for-gestational-age pregnancy and outcomes in non-diabetic fetal macrosomia. *Arch Dis Child Fetal Neonatal Ed*. 2011;96(Suppl 1):Fa89-Fa.
118. Mirkinson LJ. Hypoglycemia and Infants of Diabetic Mothers. Comprehensive Pediatric Hospital Medicine. pg 274-278. Elsevier: New York, USA 2007.
119. Monari F, Menichini D, Spano'Bascio L, et al. A First Trimester Prediction Model for Large for Gestational Age Infants: a preliminary study. *BMC Pregnancy Childbirth.* 2021;21(1):654-662
120. Weissmann-Brenner A, Simchen MJ, Zilberberg E, et al. Maternal and neonatal outcomes of large for gestational age pregnancies. *Acta Obstet Gynecol Scand*. 2012;91(7):844-849.
121. Mendez-Figueroa H, Pedroza C, Chauhan SP. Large for gestational age infants and adverse outcomes among uncomplicated pregnancies at term. *Am J Perinatol*. 2017;34(7):655-662.
122. Kuciene R, Dulskiene V, Medzioniene J. Associations between high birth weight, being large for gestational age, and high blood pressure among adolescents: a cross-sectional study. *Eur J Nutr*. 2018;57(1):373-381.
123. Lei X, Zhao D, Huang L, et al. Childhood health outcomes in term, large-for-gestational-age babies with different postnatal growth patterns. *Am J Epidemiol.* 2018;187(3):507-514.
124. Tanner JM. Growth as a target-seeking function. In: Falkner F, Tanner JM. (Eds) Human Growth. Springer: Boston, USA; 1986
125. Hediger ML, Overpeck MD, Maurer KR, et al. Growth of infants and young children born small or large for gestational age: findings from the Third National Health and Nutrition Examination Survey. *Arch Pediatr Adolesc Med.* 1998;152(12):1225-1231.

126. Pinar MH, Gibbins K, He M, Kostadinov S, Silver R. Early pregnancy losses: review of nomenclature, histopathology, and possible etiologies. *Fetal Pediatr Pathol*. 2018;37(3):191-209.
127. Jauniaux E, Johns J, Burton G. The role of ultrasound imaging in diagnosing and investigating early pregnancy failure. *Ultrasound Obstet Gynecol.* 2005;25(6):613-624.
128. Jurkovic D, Overton C, Bender-Atik R. Diagnosis and management of first trimester miscarriage. *BMJ*. 2013;346:f3676
129. Branch DW. Thoughts on the mechanism of pregnancy loss associated with the antiphospholipid syndrome. *Lupus*. 1994;3(4):275-280.
130. Farquharson RG, Jauniaux E, Exalto N, ESHRE Special Interest Group for Early Pregnancy (SIGEP). Updated and revised nomenclature for description of early pregnancy events. *Hum Reprod*. 2005;20(11):3008-3011.
131. Lockwood CJ. Prediction of pregnancy loss. *Lancet.* 2000;355(9212):1292-1293.
132. Macklon NS, Geraedts JP, Fauser BC. Conception to ongoing pregnancy: the 'black box'of early pregnancy loss. *Hum Reprod Update.* 2002;8(4):333-343.
133. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early loss of pregnancy. *N Engl J Med.* 1988;319(4):189-194
134. Semprini AE, Simoni G. Not so inefficient reproduction. *Lancet.* 2000;356(9225):257-258.
135. Lidegaard Ø, Mikkelsen AP, Egerup P, et al. Pregnancy loss: A 40-year nationwide assessment. *Acta Obstet Gynecol Scand*.2020;99(11):1492-1496.
136. Kolte AM, Bernardi LA, Christiansen OB, et al. Terminology for pregnancy loss prior to viability: a consensus statement from the ESHRE early pregnancy special interest group. *Hum Reprod.* 2015;30(3):495-498.
137. Boyraz G, Bozdağ G. Pregnancy of unknown location*. J Turk Ger Gynecol Assoc.* 2013;14(2):104-108.
138. Pereira PP, Cabar FR, Gomez ÚT, Francisco RPV. Pregnancy of unknown location. *Clinics.* 2019;74:1-7.
139. Hajenius PJ, Mol F, Mol BWJ, et al. Interventions for tubal ectopic pregnancy. *Cochrane Database Syst Rev.* 2007;2007(1):CD000324.
140. Quenby S, Gallos ID, Dhillon-Smith RK, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. *Lancet*. 2021;397(10285):1658-1667.
141. MacDorman MF, Munson ML, Kirmeyer S. Fetal and perinatal mortality, United States, 2004. *Natl Vital Stat Rep.* 2007;56(3):1-19.
142. Gourbin G, Masuy-Stroobant G. Registration of vital data: are live births and stillbirths comparable all over Europe? *Bull World Health Organ*. 1995;73(4):449-460.

143. Giakoumelou S, Wheelhouse N, Cuschieri K, et al. The role of infection in miscarriage. *Hum Reprod Update.* 2016;22(1):116-133.
144. RCOG. Management of early pregnancy loss. Green-top Guideline No. 25. London, UK: Royal College of Obstetricians and Gynaecologists; 2006.
145. Brown S. Miscarriage and its associations. *Semin Reprod Med.* 2008;26(5):391-400.
146. Regan L, Rai R. Epidemiology and the medical causes of miscarriage. *Baillieres Best Pract Res Clin Obstet Gynaecol.* 2000;14(5):839-854.
147. Christiansen OB, Elson J, Kolte AM, et al. ESHRE guideline: recurrent pregnancy loss. *Hum Reprod Open*. 2018;2018(2):hoy004-hoy.
148. Laisk T, Soares ALG, Ferreira T, et al. The genetic architecture of sporadic and multiple consecutive miscarriage. *Nat Commun.* 2020;11(1):5980-5992
149. Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. *BMC Med.* 2013;11(1):154-164.
150. Nguyen RH, Wilcox AJ. Terms in reproductive and perinatal epidemiology: 2. Perinatal terms. *J Epidemiol Community Health*. 2005;59(12):1019-1021.
151. WHO. Neonatal and perinatal mortality: country, regional and global estimates: World Health Organization; 2006.
152. Mohangoo AD, Blondel B, Gissler M, et al. International comparisons of fetal and neonatal mortality rates in high-income countries: should exclusion thresholds be based on birth weight or gestational age? *PLoS One*. 2013;8(5):e64869.
153. Lawn JE, Yakoob MY, Haws RA, et al. 3.2 million stillbirths: epidemiology and overview of the evidence review. *BMC Pregnancy Childbirth*. 2009;9(Suppl 1):S2.
154. WHO. Stillbirths. In: World Health Organization (Ed). Maternal, newborn, child and adolescent health. Geneva, Switzerland; 2017.
155. Morisaki N, Ganchimeg T, Vogel J, et al. Impact of stillbirths on international comparisons of preterm birth rates: a secondary analysis of the WHO multi-country survey of Maternal and Newborn Health. *Br J Obstet Gynaecol.* 2017;124(9):1346-1354.
156. Smith GC, Fretts RC. Stillbirth. *Lancet.* 2007;370(9600):1715-1725.
157. Goldenberg RL, Thompson C. The infectious origins of stillbirth. *Am J Obstet Gynecol.* 2003;189(3):861-873.
158. Goldenberg RL, Kirby R, Culhane J. Stillbirth: a review. *J Matern Fetal Neonatal Med.* 2004;16(2):79-94.
159. Graham N, Heazell AE. When the Fetus Goes Still and the Birth Is Tragic: The Role of the Placenta in Stillbirths. *Obstet Gynecol Clin North Am.* 2020;47(1):183-196.
160. Fretts RC, Boyd ME, Usher RH, Usher HA. The changing pattern of fetal death, 1961-1988. *Obstet Gynaecol*. 1992;79(1):35-39.

161. Reddy UM, Goldenberg R, Silver R, et al. Stillbirth classification—developing an international consensus for research: executive summary of a National Institute of Child Health and Human Development workshop. *Obstet Gynaecol*. 2009;114(4):901-914
162. Fretts RC. Etiology and prevention of stillbirth. *Am J Obstet Gynecol.* 2005;193(6):1923-1935.
163. Frøen JF, Gardosi JO, Thurmann A, Francis A, Stray-Pedersen B. Restricted fetal growth in sudden intrauterine unexplained death. *Acta Obstet Gynecol Scand*. 2004;83(9):801-807.
164. Khong TY, Mooney EE, Ariel I, et al. Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement. *Arch Pathol Lab Med*. 2016;140(7):698-713.
165. Bar J, Schreiber L, Ben-Haroush A, et al. The placental vascular component in early and late intrauterine fetal death*. Thromb Res*. 2012;130(6):901-905.
166. Man J, Hutchinson J, Heazell A, et al. Stillbirth and intrauterine fetal death: role of routine histopathological placental findings to determine cause of death. *Ultrasound Obstet Gynecol.* 2016;48(5):579-584.
167. Wu JN, Ren YY, Zhu C,et al. Abnormal placental perfusion and the risk of stillbirth: a hospital-based retrospective cohort study. *BMC Pregnancy Childbirth*. 2021;21(1):1-10.
168. Kulkarni VG, Sunilkumar KB, Nagaraj T, et al. Maternal and Fetal Vascular Lesions of Malperfusion in the Placentas Associated with Fetal and Neonatal Death: Results of a Prospective Observational Study. *Am J Obstet Gynecol.* 2021; 225(6):660.e1-660.e12.
169. Tiwari P, Gupta MM, Jain SL. Placental findings in singleton stillbirths: a case-control study from a tertiary-care center in India. *J Perinat Med.* 2021;65(1):22-25
170. Messerlian C, Basso O. Cohort studies in the context of obstetric and gynecologic research: a methodologic overview. *Acta Obstet Gynecol Scand*. 2018;97(4):371-379.
171. Petitti DB. Perinatal epidemiology: studying the effects of illness and medications during pregnancy. *Immunol Allergy Clin North Am*. 2000;20(4):673-685.
172. Giannakou K. Perinatal epidemiology: Issues, challenges, and potential solutions. *Obstet Med.* 2020; 14(2):77-82.
173. Weinberg CR, Wilcox AJ. Methodologic Issues in Reproductive Epidemiology. In: Rothman KJ GS, Lash TL (Eds). Modern Epidemiology. 3rd ed. p. 585–608.: Lippincott Williams & Wilkins: Philadelphia, PA, USA. 2008.
174. Olsen J, Basso O. Reproductive epidemiology. In: Ahrens W PI (Ed) Handbook of Epidemiology. pg 1043-111. Springer Verlag: Berlin, Germany; 2005.
175. Jackson D, Lang J, Ganiats T. Epidemiological issues in perinatal outcomes research. *Paediatr Perinat Epidemiol.* 1999;13(4):392-404.
176. Basso O, Weinberg CR, Baird DD, Wilcox AJ, Olsen J. Subfecundity as a correlate of preeclampsia: a study within the Danish National Birth Cohort. *Am J Epidemiol.* 2003;157(3):195-202.

177. Paneth N. Invited commentary: the hidden population in perinatal epidemiology. *Am J Epidemiol.* 2008;167(7):793-796.
178. Hertz-Picciotto I, Pastore LM, Beaumont JJ. Timing and patterns of exposures during pregnancy and their implications for study methods. *Am J Epidemiol.* 1996;143(6):597-607.
179. Selevan SG, Lemasters GK. The dose-response fallacy in human reproductive studies of toxic exposures. *J Occup Med*. 1987;29(5):451-454.
180. Zhang LP, Lei XP, Luo LJ, Dong WB. Risk factors for necrotizing enterocolitis in very preterm infants: a case–control study in southwest China. *J Matern Fetal Neonatal Med.* 2019;32(6):896-901.
181. Darrow LA. Invited commentary: application of case-crossover methods to investigate triggers of preterm birth. *Am J Epidemiol.* 2010;172(10):1118-1120.
182. Hernández-Díaz S, Boeke CE, Romans AT, et al. Triggers of spontaneous preterm delivery–why today? *Paediatr Perinat Epidemiol.* 2014;28(2):79-87.
183. Sagiv SK, Mendola P, Loomis D, et al. A time series analysis of air pollution and preterm birth in Pennsylvania, 1997–2001. *Environ Health Perspect*. 2005;113(5):602-606.
184. Gronlund CJ, Yang AJ, Conlon KC, et al. Time series analysis of total and direct associations between high temperatures and preterm births in Detroit, Michigan. *BMJ Open*. 2020;10(2):e032476.
185. Savitz DA, Terry JW Jr, Dole N, et al Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol.* 2002;187(6):1660-1666.
186. Martin JA. United States vital statistics and the measurement of gestational age. *Paediatr Perinat Epidemiol.* 2007;21(Suppl 2):13-21.
187. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol.* 2007;21(s2):86-96.
188. Saito M, Yazawa K, Hashiguchi A, et al. Time of ovulation and prolonged pregnancy. *Am J Obstet Gynecol.* 1972;112(1):31-38.
189. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol.* 1996;174(1 Pt 1):278-281.
190. Gardosi J, Francis A. Controlled trial of fundal height measurement plotted on customised antenatal growth charts. *Br J Obstet Gynaecol.* 1999 Apr;106(4):309-317.
191. Sunjoh F, Njamnshi AK, Tietche F, Kago I. Assessment of gestational age in the Cameroonian newborn infant: a comparison of four scoring methods. *J Trop Pediatr.* 2004;50(5):285-291.
192. Alexander GR, de Caunes F, Hulsey TC, Tompkins ME, Allen M. Validity of postnatal assessments of gestational age: a comparison of the method of Ballard et al. and early ultrasonography. *Am J Obstet Gynecol.* 1992;166(3):891-895.

193. Alexander GR, Tompkins ME, Petersen DJ, Hulsey TC, Mor J. Discordance between LMP-based and clinically estimated gestational age: implications for research, programs, and policy. *Public Health Rep.* 1995;110(4):395-402.

194. Snowden JM, Basso O. Causal inference in studies of preterm babies: a simulation study. *Br J Obstet Gynaecol.* 2018;125(6):686-692.

195. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol.* 2002;31(1):163-165.

196. Greenland SJ. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology.* 2003;14(3):300-306.

197. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol.* 2010;39(2):417-420.

198. Greenland SJ, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10(1):37-48

199. Royston P, Altman D. Design and analysis of longitudinal studies of fetal size. pg 307-12. Wiley Online Library; 1995.

200. Hansen CA, Barnett AG, Pritchard G. The effect of ambient air pollution during early pregnancy on fetal ultrasonic measurements during mid-pregnancy. *Environ Health Perspect.* 2008;116(3):362-369.

201. Gurrin LC, Blake KV, Evans SF, Newnham JP. Statistical measures of foetal growth using linear mixed models applied to the foetal origins hypothesis. *Stat Med.* 2001;20(22):3391-3409.

202. Romero R, Deter R. Should serial fetal biometry be used in all pregnancies? *Lancet*. 2015;386(10008):2038-2040.

203. Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *Lancet.* 2015;386(10008):2089-2097.

204. van den Hooven EH, Pierik FH, de Kluizenaar Y, et al. Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: a prospective cohort study. *Environ Health Perspect.* 2012;120(1):150-156.

205. Smarr MM, Vadillo-Ortega F, Castillo-Castrejon M, O'Neill MS. The use of ultrasound measurements in environmental epidemiological studies of air pollution and fetal growth. *Curr Opin Pediatr.* 2013;25(2):240-246

206. Norris T, Johnson W, Petherick E, et al. Investigating the relationship between fetal growth and academic attainment: secondary analysis of the Born in Bradford (BiB) cohort. *Int J Epidemiol.* 2018;47(5):1475-1484.

207. Woodruff TJ, Parker JD, Darrow LA, et al. Methodological issues in studies of air pollution and reproductive health. *Environ Res.* 2009;109(3):311-320.
208. Hemachandra AH, Klebanoff MA, Duggan AK, Hardy JB, Furth SL. The association between intrauterine growth restriction in the full-term infant and high blood pressure at age 7 years: results from the Collaborative Perinatal Project. *Int J Epidemiol.* 2006;35(4):871-877.
209. Hertz-Picciotto I, Swan SH, Neutra RR, Samuels SJ. Spontaneous abortions in relation to consumption of tap water: an application of methods from survival analysis to a pregnancy follow-up study. *Am J Epidemiol*. 1989;130(1):79-93.
210. Swan SH, Waller K, Hopkins B, et al. A prospective study of spontaneous abortion: relation to amount and source of drinking water consumed in early pregnancy. *Epidemiology*. 1998;9(2):126-133.
211. Harlap S. A life table of spontaneous abortions and the effects of age, parity and other variables. *J Med Genet.* 1980;1:145-58.
212. Shapiro S. Factors associated with early and late fetal loss. *Adv Planned Parent.* 1971;6:45-63.
213. Wilcox AJ, Gladen BC. Spontaneous abortion: the role of heterogeneous risk and selective fertility. *Early Hum Dev*. 1982;7(2):165-178.
214. Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. *Am J Epidemiol*. 2007;165(4):444-452.
215. Stegherr R, Allignol A, Meister R, Schaefer C, Beyersmann J. Estimating cumulative incidence functions in competing risks data with dependent left-truncation. *Stat Med.* 2020;39(4):481-493.
216. Howards PP, Schisterman EF, Heagerty PJ. Potential confounding by exposure history and prior outcomes: an example from perinatal epidemiology. *Epidemiology.* 2007; 18(5):544-51.
217. Howards PP, Schisterman EF, Poole C, Kaufman JS, Weinberg CR. "Toward a clearer definition of confounding" revisited with directed acyclic graphs. *Am J Epidemiol*. 2012;176(6):506-511.
218. Wilcox AJ, Baird DD, Weinberg CR. Time of implantation of the conceptus and loss of pregnancy. *N Engl J Med.* 1999;340(23):1796-1799.
219. Venners SA, Korrick S, Xu X, et al. Preconception serum DDT and pregnancy loss: a prospective study using a biomarker of pregnancy. *Am J Epidemiol.* 2005;162(8):709-716.
220. Weinberg CR, Hertz-Picciotto I, Baird DD, Wilcox AJ. Efficiency and bias in studies of early pregnancy loss. *Epidemiology.* 1992;3(1):17-22.
221. Knez J, Day A, Jurkovic D. Ultrasound imaging in the management of bleeding and pain in early pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2014;28(5):621-636.
222. Wieringa-de Waard M, Bonsel GJ, Ankum WM, Vos J, Bindels PJ. Threatened miscarriage in general practice: diagnostic value of history taking and physical examination. *Br J Gen Pract.* 2002;52(483):825-829.

223. National Collaborating Centre for Women's and Children's Health (UK). Ectopic Pregnancy and Miscarriage: Diagnosis and Initial Management in Early Pregnancy of Ectopic Pregnancy and Miscarriage. RCOG Press: London, UK; 2012.
224. Kirk E, Condous G, Van Calster B, et al. Rationalizing the follow-up of pregnancies of unknown location. *Hum Reprod.* 2007;22(6):1744-1750.
225. Dickey RP, Olar TT, Taylor SN, Curole DN, Matulich EM. Relationship of small gestational sac-crown-rump length differences to abortion and abortus karyotypes. *Obstet Gynaecol*. 1992;79(4):554-557.
226. Laboda L, Estroff J, Benacerraf B. First trimester bradycardia. A sign of impending fetal loss. *J Ultrasound Med*. 1989;8(10):561-563.
227. Norwitz ER, Schorge JO. Obstetrics and Gynecology at a Glance. John Wiley & Sons: New Jersey, USA 2013.
228. Lindsey JL. Evaluation of fetal death. In: Witlin A, Talavera F, Smith CV, Gaupp FB, Shulman LP (Eds) *eMedicine.* Retrieved from https://emedicine.medscape.com/article/259165-overview Last updated January 24, 2022
229. Thompson C. Fortuitous Phenomena: On Complexity, Pragmatic Randomised Controlled Trials, and Knowledge for Evidence-Based Practice*. Worldviews Evid Based Nurs.* 2004;1(1):9-17.
230. Tooher RL, Middleton PF, Crowther CA. A thematic analysis of factors influencing recruitment to maternal and perinatal trials. *BMC Pregnancy Childbirth*. 2008;8(1):1-12.
231. Faraoni D, Schaefer ST. Randomized controlled trials vs. observational studies: why not just live together? *BMC Anesthesiol.* 2016;16(1):1-4.
232. Rosenbaum PR. Observational studies. Observational studies: Springer; 2002. p. 1-17.
233. Kumar G, Acharya AS. Biases in epidemiological studies: How far are we from the truth? *India J Med Spec.* 2014;5(1):29-35.
234. Sackett DL. Bias in analytic research. The case-control study consensus and controversy. pg 51-63. Elsevier: New York, USA; 1979.
235. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. *Nephron Clin Pract.* 2010;115(2):c94-c9.
236. Kesmodel US. Information bias in epidemiological studies with a special focus on obstetrics and gynecology. *Acta Obstet Gynecol Scand*. 2018;97(4):417-423.
237. Dunne J, Tessema GA, Ognjenovic M, Pereira G. Quantifying the influence of bias in reproductive and perinatal epidemiology through simulation. *Ann Epidemiol*. 2021;63:86-101.
238. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Springer Science & Business Media: New York, USA; 2011.
239. Chen H, Cohen P, Chen S. Biased odds ratios from dichotomization of age. *Stat Med*. 2007;26(18):3487-3497.

240. Waller DK SW, Gu Y, Cunningham GC. Assessing number-specific error in the recall of onset of last menstrual period. *Paediatr Perinat Epidemiol*. 2000;14(3):263–267.
241. Tentoni S, Astolfi P, De Pasquale A, Zonta LA. Birthweight by gestational age in preterm babies according to a Gaussian mixture model. *Br J Obstet Gynaecol*. 2004;111(1):31-37.
242. Gjessing HK, Skjaerven R, Wilcox AJ. Errors in gestational age: evidence of bleeding early in pregnancy. *Am J Public Health*. 1999;89(2):213-218.
243. Morin I, Morin L, Zhang X, et al. Determinants and consequences of discrepancies in menstrual and ultrasonographic gestational age estimates. *Br J Obstet Gynaecol*. 2005;112(2):145-152.
244. Brakenhoff TB, Mitroiu M, Keogh RH, et al. Measurement error is often neglected in medical literature: a systematic review. *J Clin Epidemiol.* 2018;98:89-97.
245. Gustafson P. Measurement error and misclassification in statistics and epidemiology: impacts and Bayesian adjustments. CRC Press: Florida, USA; 2003.
246. Carroll RJ, Ruppert D, Crainiceanu CM, Stefanski LA. Measurement error in nonlinear models: a modern perspective. Chapman and Hall/CRC: Florida, USA 2006.
247. Fuller WA. Measurement error models. John Wiley & Sons: New Jersey, USA; 2009.
248. Buonaccorsi JP. Measurement error: models, methods, and applications. Chapman and Hall/CRC: Florida, USA; 2010.
249. Stefanski LA. Measurement error models*. J Am Stat Assoc.* 2000;95(452):1353-1358.
250. Guolo A. Robust techniques for measurement error correction: a review. *Stat Methods Med Res.* 2008;17(6):555-580.
251. Keogh RH, White IR. A toolkit for measurement error correction, with a focus on nutritional epidemiology. *Stat Med*. 2014;33(12):2137-2155.
252. Dudley NJ. The management of error in ultrasound fetal growth monitoring. *Ultrasound.* 2021;29(1):4-9.
253. Källén K. Mid-trimester ultrasound prediction of gestational age: advantages and systematic errors. *Ultrasound Obstet Gynecol*. 2002;20(6):558-563.
254. Grace YY. Statistical analysis with measurement error or misclassification: Springer: New York, USA; 2016.
255. Rothman KJ, Greenland S, Lash TL. Modern epidemiology: Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, USA; 2008.
256. Fuller WA. Measurement Error Models. John Wiley & Sons: New Jersey, USA 1987.
257. Bachl M, Scharkow M. Some suggestions on dealing with measurement error in linkage analyses. S*ociology.* 2017.
258. Rosner B, Willett W, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med.* 1989;8(9):1051-1069.

259. Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*. 1982;69(2):331-342.
260. Lash TL, Fox MP, MacLehose RF, et al. Good practices for quantitative bias analysis. *Int J Epidemiol.* 2014;43(6):1969-1985.
261. Blackwell M, Honaker J, King G. A unified approach to measurement error and missing data: Details and extensions*. Sociol Methodol.* 2017;46(3):342-369.
262. Blackwell M, Honaker J, King GJ. A unified approach to measurement error and missing data: overview and applications. *Sociol Methodol.* 2017;46(3):303-341.
263. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol.* 2018;47(1):226-235.
264. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. *Int J Epidemiol.* 2018;47(5):1714-1722.
265. Elwert F, Winship C. Endogenous selection bias: The problem of conditioning on a collider variable. *Annu Rev Sociol*. 2014;40:31-53.
266. Snowden JM, Bovbjerg ML, Dissanayake M, Basso O. The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection. *Curr Epidemiol Rep.* 2018;5(4):379–387.
267. Nguyen TQ, Dafoe A, Ogburn EL. The magnitude and direction of collider bias for binary variables*. Epidemiol Methods*. 2019;8(1):1-29.
268. Neophytou AM, Kioumourtzoglou MA, Goin DE, Darwin KC, Casey JA. Educational note: addressing special cases of bias that frequently occur in perinatal epidemiology. *Int J Epidemiol.* 2021;50(1):337-345.
269. Shringarpure S, Xing EP. Effects of sample selection bias on the accuracy of population structure and ancestry inference. *G3 (Bethesda).* 2014;4(5):901-911.
270. Chard T, Seppala M. Frequency of Implantation and Early Pregnancy Loss in Natural Cycles. *Baillieres Clin Obstet Gynaecol*. 1991;5(1):179-189.
271. Liew Z, Olsen J, Cui X, Ritz B, Arah OA. Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants. *Int J Epidemiol.* 2015;44(1):345-354.
272. Joseph KS, Kramer MS. The fetuses-at-risk approach: survival analysis from a fetal perspective. *Acta Obstet Gynecol Scand*. 2018;97(4):454-465.
273. Raz R, Kioumourtzoglou MA, Weisskopf MG. Live-birth bias and observed associations between air pollution and autism. *Am J Epidemiol*. 2018;187(11):2292-2296.
274. Leung M, Kioumourtzoglou MA, Raz R, Weisskopf MG. Bias due to Selection on Live Births in Studies of Environmental Exposures during Pregnancy: A Simulation Study. *Environ Health Perspect.* 2021;129(4):047001.
275. Hernán MA, Robins JM. Causal Inference: What If. Chapman & Hall/CRC: Florida, USA; 2020

276. Lisonkova S, Joseph KS. Left truncation bias as a potential explanation for the protective effect of smoking on preeclampsia. *Epidemiology*. 2015;26(3):436-440.
277. Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*. 2010;340:b5087.
278. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167(4):492-499.
279. Mansournia MA, Nazemipour M, Etminan M. Causal diagrams for immortal time bias. *Int J Epidemiol.* 2021;50(5):1405-1409.
280. Hutcheon JA, Savitz DA. Invited commentary: influenza, influenza immunization, and pregnancy—it's about time. *Am J Epidemiol*. 2016;184(3):187-191.
281. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* 2004;15(5):615-625.
282. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. *Eur Respir J.* 2004;23(3):391-395.
283. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*. 2007;16(3):241-249.
284. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology.* 2009;20(4):488-495.
285. Ananth CV, Schisterman EF. Confounding, causality, and confusion: the role of intermediate variables in interpreting observational studies in obstetrics*. Am J Obstet Gynecol*. 2017;217(2):167-175.
286. VanderWeele TJ, Mumford SL, Schisterman EF. Conditioning on intermediates in perinatal epidemiology. *Epidemiology*. 2012;23(1):1-19.
287. Kiely JL, Sergievsky GH. Some conceptual problems in multivariable analyses of perinatal mortality. *Paediatr Perinat Epidemiol*. 1991;5(3):243-257.
288. Kiely JL, Kleinman JC. Birth-weight-adjusted infant mortality in evaluations of perinatal care: towards a useful summary measure. *Stat Med.* 1993;12(3-4):377-392.
289. Kramer MS. Commentary: Biology vs. methodology in investigating causal pathways for infant mortality. *Paediatr Perinat Epidemiol*. 2009;23(5):414-416.
290. Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight “paradox” uncovered? *Am J Epidemiol*. 2006;164(11):1115-1120.
291. Whitcomb BW, Schisterman EF, Perkins NJ, Platt RW. Quantification of collider-stratification bias and the birthweight paradox. *Paediatr Perinat Epidemiol.* 2009;23(5):394-402.
292. Schisterman EF, Whitcomb BW, Mumford SL, Platt RW. Z-scores and the birthweight paradox. *Paediatr Perinat Epidemiol*. 2009;23(5):403-413.
293. Wilcox AJ. Birth weight and perinatal mortality: the effect of maternal smoking. *Am J Epidemiol*. 1993;137(10):1098-1104.

294. Stokholm L, Juhl M, Talge NM, et al. Obstetric oxytocin exposure and ADHD and ASD among Danish and Finnish children. *Int J Epidemiol.* 2021;50(2):446-456.
295. Agier L, Basagaña X, Hernandez-Ferrer C, et al. Association between the pregnancy exposome and fetal growth. *Int J Epidemiol.* 2020;49(2):572-586.
296. Gao M, Scott K, Koupil I. Associations of perinatal characteristics with endometriosis: a nationwide birth cohort study. *Int J Epidemiol.* 2020;49(2):537-547.
297. Darling AM, Werler MM, Cantonwine DE, Fawzi WW, McElrath TF. Timing and amount of gestational weight gain in association with adverse birth outcomes. *Epidemiology*. 2019;30(5):695-705.
298. Sachdeva P, Patel B, Patel B. Drug use in pregnancy; a point to ponder! *Indian J Pharm Sci*. 2009;71(1):1-7.
299. Buhimschi CS, Weiner CP. Medications in pregnancy and lactation: part 1. *Teratology Obstet Gynecol.* 2009;113(1):166-188.
300. Wood ME, Lapane KL, van Gelder MM, Rai D, Nordeng HM. Making fair comparisons in pregnancy medication safety studies: An overview of advanced methods for confounding control. *Pharmacoepidemiol Drug Saf.* 2018;27(2):140-147.
301. Bérard A, Sheehy O. The Quebec Pregnancy Cohort–prevalence of medication use during gestation and pregnancy outcomes. *PLoS One*. 2014;9(4):e93870.
302. Pisa FE, Casetta A, Clagnan E, et al. Medication use during pregnancy, gestational age and date of delivery: agreement between maternal self-reports and health database information in a cohort. *BMC Pregnancy Childbirth*. 2015;15(1):1-14.
303. Artama M, Gissler M, Malm H, Ritvanen A, Drugs, Group PS. Nationwide register-based surveillance system on drugs and pregnancy in Finland 1996–2006. *Pharmacoepidemiol Drug Saf*. 2011;20(7):729-738.
304. Howard TB, Tassinari MS, Feibus KB, Mathis LL, editors. Monitoring for teratogenic signals: Pregnancy registries and surveillance methods. *Am J Med Genet Part C Semin Med Genet.* 2011;157(3):209–214.
305. Andrade SE, Raebel MA, Morse AN, et al. Use of prescription medications with a potential for fetal harm among pregnant women. *Pharmacoepidemiol Drug Saf*. 2006;15(8):546-554.
306. Andrade SE, Bérard A, Nordeng HM, et al. Administrative claims data versus augmented pregnancy data for the study of pharmaceutical treatments in pregnancy. *Curr Epidemiol Rep*. 2017;4(2):106-116.
307. Charlton RA, Neville AJ, Jordan S, et al. Healthcare databases in Europe for studying medicine use and safety during pregnancy. *Pharmacoepidemiol Drug Saf*. 2014;23(6):586-594.
308. Temmerman M, Plummer FA, Mirza NB, et al. Infection with HIV as a risk factor for adverse obstetrical outcome. *AIDS*. 1990;4(11):1087-1094.

309. Brocklehurst P, French R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *Br J Obstet Gynaecol*. 1998;105(8):836-848.

310. Taha TE, Dallabetta GA, Canner JK, et al. The effect of human immunodeficiency virus infection on birthweight, and infant and child mortality in urban Malawi. *Int J Epidemiol*. 1995;24(5):1022-1029.

311. Bulterys M, Chao A, Munyemana S, et al. Maternal human immunodeficiency virus 1 infection and intrauterine growth: a prospective cohort study in Butare, Rwanda. *Pediatr Infect Dis J*. 1994;13(2):94-99.

312. Parekh N, Ribaudo H, Souda S, et al. Risk factors for very preterm delivery and delivery of very-small-for-gestational-age infants among HIV-exposed and HIV-unexposed infants in Botswana. *Int J Gynaecol Obstet*. 2011;115(1):20-25.

313. Rollins NC, Coovadia HM, Bland RM, et al. Pregnancy outcomes in HIV-infected and uninfected women in rural and urban South Africa. *J Acquir Immune Defic Syndr*. 2007;44(3):321-328.

314. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Geneva, Switzerland: World Health Organization; 2016.

315. Townsend CL, Cortina-Borja M, Peckham CS, Tookey PA. Antiretroviral therapy and premature delivery in diagnosed HIV-infected women in the United Kingdom and Ireland. *AIDS.* 2007;21(8):1019-1026.

316. Thorne C, Patel D, Newell ML. Increased risk of adverse pregnancy outcomes in HIV-infected women treated with highly active antiretroviral therapy in Europe. *AIDS*. 2004;18(17):2337-2339.

317. 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):1195-1201.

318. Szyld EG, Warley EM, Freimanis L, et al. Maternal antiretroviral drugs during pregnancy and infant low birth weight and preterm birth. *AIDS*. 2006;20(18):2345-2353.

319. Tuomala RE, Shapiro DE, Mofenson LM, et al. Antiretroviral therapy during pregnancy and the risk of an adverse outcome. *N Engl J Med*. 2002;346(24):1863-1870.

320. Tuomala RE, Watts DH, Li D, et al. Improved obstetric outcomes and few maternal toxicities are associated with antiretroviral therapy, including highly active antiretroviral therapy during pregnancy. *J Acquir Immune Defic Syndr*. 2005;38(4):449-473.

321. Boer K, Nellen JF, Patel D, et al. The AmRo study: pregnancy outcome in HIV-1-infected women under effective highly active antiretroviral therapy and a policy of vaginal delivery. *Br J Obstet Gynaecol*. 2007;114(2):148-155.

322. Martin F, and Taylor, GP. Increased rates of pre-term delivery are associated with the initiation of highly active antiretroviral therapy during pregnancy: a single centre cohort study. *J Infect Dis*. 2007;196(4):558–561.

323. Sibiude J, Warszawski J, Tubiana R, et al. Premature delivery in HIV-infected women starting protease inhibitor therapy during pregnancy: role of the ritonavir boost? *Clin Infect Dis.* 2012;54(9):1348-1360.

324. Lopez M, Figueras F, Hernandez S, et al. Association of HIV infection with spontaneous and iatrogenic preterm delivery: effect of HAART. *AIDS*. 2012;26(1):37-43.

325. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.

326. Townsend CL, Schulte J, Thorne C, et al. Antiretroviral therapy and preterm delivery — a pooled analysis of data from the United States and Europe. *Br J Obstet Gynaecol*. 2010;117(11):1399-1410.

327. Fiore S, Newell ML, Trabattoni D, et al. Antiretroviral therapy-associated modulation of Th1 and Th2 immune responses in HIV-infected pregnant women. *J Reprod Immunol*. 2006;70(1):143-150.

328. Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet.* 2010;375(9730):1969-1987.

329. Lawn JE, Wilczynska-Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. *Int J Epidemiol.* 2006;35(3):706-718.

330. Luu TM, Mian MOR, Nuyt AM. Long-term impact of preterm birth: neurodevelopmental and physical health outcomes. *Clin Perinatol*. 2017;44(2):305-314.

331. Patel RM. Short-and long-term outcomes for extremely preterm infants. *Am J Perinatol*. 2016;33(03):318-328.

332. Katz J, Lee AC, Kozuki N, Black RE. Mortality risk among term and preterm small for gestational age infants. *Nestle Nutr Inst Workshop Ser.* 2015;81:29-35

333. Blencowe H, Krasevec J, de Onis M, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *Lancet Glob Health.* 2019;7(7):e849-e60.

334. Bailey H, Zash R, Rasi V, Thorne C. HIV treatment in pregnancy. *Lancet HIV*. 2018;5(8):e457-e67.

335. Watts DH, Williams PL, Kacanek D, et al. Combination antiretroviral use and preterm birth. *J Infect Dis*. 2013;207(4):612-621.

336. Zash R, Jacobson DL, Diseko M, et al. Comparative safety of antiretroviral treatment regimens in pregnancy. *JAMA Pediatr*. 2017;171(10):e172222-e.

337. Powis KM, Kitch D, Ogwu A, et al. Increased risk of preterm delivery among HIV-infected women randomized to protease versus nucleoside reverse transcriptase inhibitor-based HAART during pregnancy *J Infect Dis*. 2011;204(4):506-514.
338. Papp E, Mohammadi H, Loutfy MR, et al. HIV protease inhibitor use during pregnancy is associated with decreased progesterone levels; a potential mechanism contributing to fetal growth restriction. *J Infect Dis*. 2015;211(1):10-18.
339. Zash R, Souda S, Chen JY, et al. Reassuring birth outcomes with tenofovir/emtricitabine/efavirenz used for prevention of mother-to-child transmission of HIV in Botswana. *J Acquir Immune Defic Syndr.* 2016;71(4):428-436.
340. WHO. Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV: interim guidelines: supplement to the 2016 consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. World Health Organization; 2018.
341. Lockman S, Brummel SS, Ziemba L, et al. Efficacy and safety of dolutegravir with emtricitabine and tenofovir alafenamide fumarate or tenofovir disoproxil fumarate, and efavirenz, emtricitabine, and tenofovir disoproxil fumarate HIV antiretroviral therapy regimens started in pregnancy (IMPAACT 2010/VESTED): a multicentre, open-label, randomised, controlled, phase 3 trial. *Lancet*. 2021;397(10281):1276-1292.
342. Uthman OA, Nachega JB, Anderson J, et al. Timing of initiation of antiretroviral therapy and adverse pregnancy outcomes: a systematic review and meta-analysis. *Lancet HIV*. 2017;4(1):e21-e30.
343. Malaba TR, Phillips TK, Le Roux S, et al. Antiretroviral therapy use during pregnancy and adverse birth outcomes in South African women. *Int J Epidemiol.* 2017;46(5):1678-1689.
344. Nlend AEN, Motazé AN, Tetang SM, et al. Preterm Birth and Low Birth Weight after In Utero Exposure to Antiretrovirals Initiated during Pregnancy in Yaoundé, Cameroon. *PLoS One*. 2016;11(3):e0150565.
345. Hoffman RM, Brummel SS, Britto P, Pilotto JH, Masheto G, Aurpibul L, et al. Adverse pregnancy outcomes among women who conceive on antiretroviral therapy. *Clin Infect Dis*. 2019;68(2):273-279.
346. Stringer J, Stoner MC, Kasaro MP, Vwalika B, Cole SR. Preconception ART and preterm birth: real effect or selection bias. *Lancet HIV*. 2017;4(4):e150.
347. Stoner MC, Cole SR, Price J, Winston J, Stringer JS. Timing of initiation of antiretroviral therapy and risk of preterm birth in studies of HIV-infected pregnant women: The role of selection bias. *Epidemiology*. 2018;29(2):224-229.
348. Phillippi JC, Neal JL, Carlson NS, et al. Utilizing datasets to advance perinatal research. *J. Midwifery Women's Health*. 2017;62(5):545-561.

349. Loken E, Gelman AJ. Measurement error and the replication crisis. *Science.* 2017;355(6325):584-585.
350. Brakenhoff TB, van Smeden M, Visseren FL, Groenwold RHH. Random measurement error: Why worry? An example of cardiovascular risk factors. *PLoS One.* 2018;13(2):e0192298.
351. Bengtson AM, Westreich D, Musonda P, et al. Multiple overimputation to address missing data and measurement error: application to HIV treatment during pregnancy and pregnancy outcomes. *Epidemiology*. 2016;27(5):642-650.
352. Westreich D. Berkson's bias, selection bias, and missing data. *Epidemiology*. 2012;23(1):159-164.
353. Arnold BF, Ercumen A, Benjamin-Chung J, Colford JM Jr. Brief report: negative controls to detect selection bias and measurement bias in epidemiologic studies. *Epidemiology.* 2016;27(5):637-641.

Chapter 3: Methodological considerations for preterm birth research

Malaba TR, Newell ML, Myer L, Ramokolo V. Methodological considerations for preterm birth research. *Front Glob Womens Health*. 2022;2:821064.

Relevance of this paper to the thesis

This chapter provides a review that expands on the methodological issues highlighted in the literature review chapter of this thesis and highlights the important selection and measurement bias issues that need to be considered for population-level estimations and epidemiological studies of preterm birth and its sequelae. Many etiological pathways for preterm birth exist, however this paper focuses on maternal infections as this is the most common pathway in low-to-middle income countries. Given the relationship between maternal HIV infection, antiretroviral therapy (ART) and adverse birth outcomes in sub-Saharan Africa, this paper further illustrates the impact of specific selection and measurement biases on the findings and interpretations of previous studies investigating the relationship between HIV/ART and preterm births. This chapter addresses Objective 1, and based on the findings of this review, subsequent analyses in this thesis attempted to mitigate the impact of these biases where possible or consider during interpretation as possible explanations for observed findings.

Contribution of the student and co-authors

TRM conceptualised the perspective presented in this review and wrote the first draft of the manuscript, under the guidance of VR. VR and MLN revised the manuscript and interpretation of findings. All authors reviewed and approved the final manuscript.

3.1 Abstract

Complications from preterm birth are a leading cause of infant mortality, with long-term implications for morbidity and quality of life of preterm infants. There are many important risk factors for preterm births however in this article, we focus on the maternal infection etiological pathway, given its significance in low-to-middle income countries. In high preterm birth settings such as sub-Saharan Africa, maternal HIV infection and antiretroviral therapy (ART) use have been associated with an increased risk of preterm births. Consequently, we highlight methodological considerations related to selection and measurement bias in preterm birth research. We further illustrate the potential impact of these biases in studies investigating the relationship between HIV/ART and preterm births. We also briefly discuss issues related to population-level estimations based on routinely collected clinical or civil registration data. We conclude by emphasising the importance of strengthening of antenatal care services to improve quality of population data as well as optimising current and future study designs, by taking into account the important methodological considerations described in this article.

3.2 Introduction

Preterm birth (PTB) is defined by the duration of gestation at the time of delivery with a cut-off of 37 weeks distinguishing preterm from term infants. As the leading cause of neonatal and child mortality and its impact on child development it has possible long-term implications for quality of life. An estimated 14.8 million PTB occur globally each year (1), however they are not equally distributed. Low-to-middle income countries (LMICs) have a higher burden than other countries; in particular sub-Saharan Africa and South Asia which account for 60-80% of global PTB (1, 2). Addressing PTB in LMICs is therefore essential for accelerating progress towards achieving Sustainable Development Goal (SDG) 3 “Ensuring healthy lives and wellbeing for all at all ages” partly by ending preventable deaths of newborns and children under five by 2030 (3). Continued focus on improving the survival and quality of life of these infants born too soon is needed - with an emphasis on optimising appropriate evidence-based PTB prevention and mitigation interventions in LMICs. High-quality epidemiologic data are necessary to determine current and emergent modifiable risk factors during the preconception, pregnancy and peripartum period. This is best achieved through well-designed studies which take into account important methodological considerations. In this article we briefly discuss some of these considerations.

3.3 Epidemiologic study of preterm in low-to-middle income countries

Preterm birth is a complex syndrome with varying phenotypes and multifactorial aetiologies, each with distinct biological pathways. In addition to gestational age classification, PTB can also be classified phenotypically according to clinical presentation into spontaneous or medically indicated PTB (4). At the individual level accurate PTB identification is important for appropriate clinical management, while at the population level it is crucial for informing policy formulation and resource allocation (5).

The epidemiologic study of PTB requires robust methods of quantifying population-level estimates, through national civil registration and vital statistics (CRVS) systems that report PTB (1). However, many LMICs still have inadequate health information systems and sub-standard statistical capacity (6). This is demonstrated by the fact that the regions with the highest PTB burden, sub-Saharan Africa and South Asia, contributed little to no CRVS data in the latest global PTB estimates (1). Most data used to estimate PTB in these settings is not nationally representative. Instead, the data come from surveys (e.g. Demographic and Health Surveys), poorly designed small-scale research studies and/or secondary and tertiary care facilities which cater to specific subsets of the total population. Consequently, in these settings, high-quality real-time data from prospective studies and secondary analyses of robust existing data are critical. This is important for standard analyses as well as novel artificial intelligence techniques like machine learning (ML), which are being increasingly used. The best ML models use high quality real time and existing data for predictive modelling and early diagnosis of health outcomes such as PTB (7). Unlike traditional statistical models, ML can handle complex data structures. Furthermore their ability to incorporate different types of data (e.g. laboratory tests, imaging and clinical notes) can provide better outcome phenotyping (8).

Pathways to spontaneous PTB include decidual haemorrhage, maternal/fetal hypothalamic-pituitary-adrenal activation, pathologic uterine overdistension or cervical) and infection/inflammation the most common pathway in LMICs (9, 10). Accordingly, understanding the overall contribution of infections (and their treatment) to PTB is essential. Given that PTB is a common outcome, small increases in risk can have substantial public health impact in settings where neonatal care services are limited. This is particularly important in sub-Saharan Africa, the epicentre of the HIV infection pandemic (11) and where HIV is one of the leading complications of pregnancy. It should be noted that antiretroviral therapy (ART) for HIV treatment, while essential for improving maternal health and survival, has also been implicated in increased PTB risk (12-14).

3.4 Methodological Considerations

In cases where the practical and ethical costs of conducting randomised trials in pregnant women are high, observational study designs are desirable, however they are subject to a variety of potential biases (15). When considering study design, data collection and analytical procedures for these studies, taking selection and information bias into account is critical. Preterm birth analyses are further subject to bias because they are not usually the primary objective of the study but are rather planned or unplanned secondary analyses. Consequently data collection procedures are often not designed to examine important aspects of PTB, so the best high-quality measures and procedures are not always utilised. To strengthen the validity of associations drawn from these observational studies of PTB, it is important to identify and evaluate potential sources of bias – especially those common and unique to perinatal research.

3.4.1 Maternal antiretroviral use and preterm birth: methodological example

In LMICs numerous studies have investigated the infection pathway to PTB, given its contribution to PTB in these settings. As an example, we draw on the vast body of literature on the relationship between HIV/ART and preterm birth in high HIV burden settings. We also highlight some methodological nuances that underlie these epidemiologic investigations. It should be noted that these concepts can be extended to other infections and/or exposures.

The relationship between maternal ART use and PTB has been an area of research for many years, with conflicting results from multiple studies. HIV infection has been shown to increase the risk of PTB through poor maternal health, increased risk factors for coinfections and fetal HIV infection(16). ART use introduces complexity to deciphering associations with PTB and other adverse birth outcomes, because it improves maternal health and reduces acute retroviral fetal infection. It would seem logical that healthier women have better birth outcomes; however, ART could also increase adverse birth outcomes through other mechanisms. This complicates the understanding of these competing forces and makes the epidemiology of this challenging.

In studies of among women receiving ART in pregnancy, some have suggested increased risk of PTB while others have found no evidence of associations with PTB. Putative explanations for these inconsistent findings have been linked to differences in study designs, study populations, HIV treatment guidelines and analytical approaches. The exposure (maternal ART use) and outcome (PTB) are both significantly impacted by selection and measurement bias (Figure 3.1).

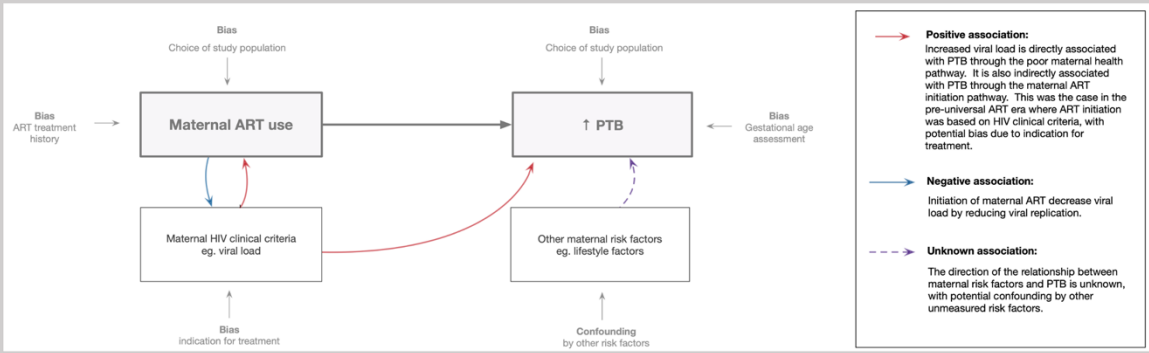

Figure 3.1 Methodological perspective: association between maternal ART use and preterm birth

Maternal ART use assessment is influenced by treatment history (including timing of initiation and ART regimens used) and study population chosen. Preterm birth assessment is influenced by the accuracy of gestational age measurement and the study population (i.e. trimester of inclusion of women in the study)

Using examples of recent studies, we highlight instances where selection and measurement bias impacted study findings.

3.4.1.1 Selection Bias

This is the distortion of an association due to the omission or inclusion of specific groups of women, such that the sample no longer reflects the population of interest (17). Perinatal study populations are dynamic and complex because the reproductive process spans from fertilisation and implantation to clinically recognised pregnancy, and further to birth and early childhood. Processes of selection (e.g. implantation failure, early pregnancy losses) and attrition (e.g. stillbirths, neonatal deaths) render these populations incompletely observable (18, 19). Studies investigating outcomes of pregnancy are subject to further selection, because recruitment is often based on a convenient sample of women willing and able to access routine care services early in pregnancy. Given that early antenatal care initiation is sub-optimal in LMICs (20), the women recruited likely represent those with better health seeking behaviours. During enrolment, additional selection can occur if women do not meet study specific eligibility criteria. This is often related to gestational age and/or the absence of an exposure or outcome of interest. Clinical interventions delivered during pregnancy (e.g. elective terminations, activity restriction or induction) can also directly influence PTB incidence and impact study findings.

Inappropriate analytical decisions can also introduce selection bias and impact the association between an exposure and PTB. For example, only including live born preterm infants, as often happens in such analyses, essentially selects for pregnancy loss. However, if an unmeasured confounder is associated with both pregnancy loss and preterm birth (Figure 3.2A), a non-causal

pathway will be created potentially biasing estimates (19). Selection bias can also be induced through inappropriate variable treatment. Because PTB pathology is incompletely understood and unmeasured confounders cannot be adjusted for, it is important to ensure independent risk factors are not treated as potential confounders, as they could potentially be intermediate variables. For example, previous PTB is often adjusted for because of its strong ties with future PTB risk (21). However, if unmeasured confounders are associated with both the intermediate variable and PTB, then this adjustment will create a non-causal pathway and potentially bias estimates (Figure 3.2B)(19).

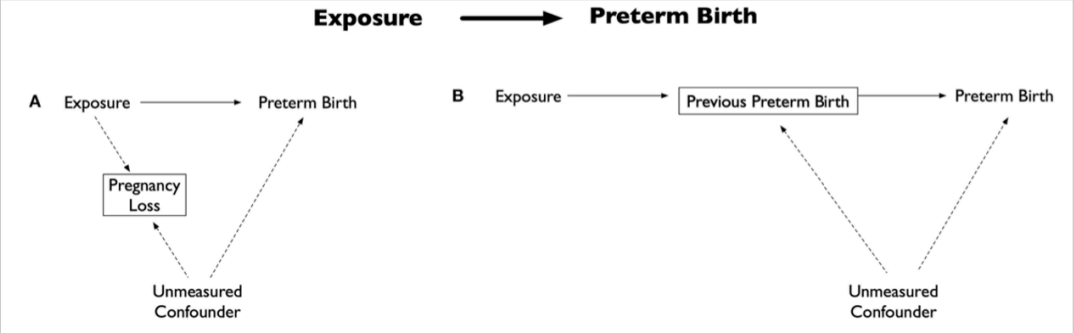

Figure 3.2 Directed acyclic graph illustrating selection bias examples

Directed acyclic graph illustrating selection bias examples. **(A)** Potential bias by inclusion of only live births. The exposure is associated with the outcome (preterm birth) and pregnancy loss, while the outcome and pregnancy loss are associated with an unmeasured confounder and preterm birth. **(B)** Potential bias by adjusting for an intermediate variable. Commonly adjusted intermediate variable for the exposure and preterm birth is preterm birth. Adjustment of previous preterm birth results in creation of a non-causal pathway between the exposure, intermediate variable, unmeasured confounder and preterm birth.

Adapted from: Neophytou AM, Kioumourtzoglou M-A, Goin DE, Darwin KC, Casey JA. Addressing special cases of bias that frequently occur in perinatal epidemiology. *Int J Epidemiol*. 2021;50(1):337-345.

Selection Bias: HIV/ART and preterm birth examples

An analysis of birth outcomes including PTB was conducted in Malawi (22) among healthy women living with and without HIV. The authors hypothesised that ART use in pregnancy would eliminate previously observed differences in pregnancy outcomes by HIV status. Study enrolment occurred at delivery, with GA assessed postnatally (Ballard score). The overall PTB rate was 10%, with no differences observed by HIV status or ART status (initiation before or during pregnancy). The authors concluded that their results were consistent with their hypothesis that the efavirenz-based ART regimen eliminates differences previously attributed to HIV infection without introducing additional risk. However, these results differ from recent studies using similar ART regimens which found an increased risk of PTB by HIV/ART status (12, 14, 23). Although the authors dismissed the possibility of selection bias, there is a potential concern around the puerperium recruitment from a few facilities. This strategy likely excluded women with serious pregnancy complications, preterm infants, acutely ill neonates and/or neonatal deaths (24). The absence of very preterm (<32 weeks) and very low birth weight (<1500 grams) infants in the study provides evidence for this. In additional to This unrepresentative study population could have led to bias towards the null resulting in the lack of difference observed by HIV and ART status (24).

Another example can be found in a systematic review conducted among women living with HIV initiating ART before or during pregnancy, at a time when few women in LMICs conceived on ART (25). Preterm birth, reported in 10 studies, showed an increased risk among preconception initiators. It was hypothesised to be due to confounding by indication, with women on preconception ART likely to be at a more advanced HIV stage. The authors acknowledged the possibility of selection bias, emanating from the unequal opportunity for PTB between women starting late in pregnancy compared to those initiating ART preconception or early. Based on the selection concerns of this systematic review, a simulation trial was conducted to quantify the impact of the exclusion of women delivering pre-ART initiation, which shows that this exclusion lowered the risk of women included in studies investigating this association (26). In this simulation women were "recruited" preconception and randomised 1:1 to immediate ART (i.e. preconception) or delayed ART (i.e. during pregnancy). Gestational age at ART initiation and at delivery were based on previously collected Zambian data (27). Preterm birth rates were compared in intention-to-treat (ITT) and naïve analyses. The ITT analysis compared all randomised women by ART initiation timing. While the naïve analysis, mimicking previous observational studies, compared all women randomised to immediate ART to a subset of women randomised to delayed ART (excluding those delivering pre-ART initiation). The ITT analysis showed no association with preconception initiation, while the naïve analysis showed an increased risk with preconception initiation.

The selection biases highlighted here apply to all analyses investigating PTB and time-dependent exposures, because there will always be exclusion of women who deliver before experiencing or receiving the exposure later in pregnancy (26).

3.4.1.2 Information Bias

A certain amount of error is intrinsic in any measurement process. Information bias is the distortion of an association caused by the inaccurate measurement of key variables (28). This bias arises when information is either not accurately collected/measured or there is discord between study definitions and true definitions, leading to exposure or outcome misclassification (29). Gestational age (GA) is central in PTB studies; however it can only be estimated approximately because fertilisation is a silent event. All methods of GA assessment (last menstrual period (LMP), symphysis fundal height measurement (SFH), and ultrasonography) are prone to varying degrees of error, primarily based on timing of assessment during pregnancy (30-32). Ultrasound is generally considered most reliable; however it is less accurate if carried out >24 weeks of pregnancy (32). SFH assessment is difficult early in pregnancy (33), and LMP reliability is limited by irregular menstrual cycles and inaccurate recall of dates (31). Postnatal assessments of the newborn by physical

examination and neuromuscular assessment are even less precise (e.g. Ballard score)(34). Use of imprecise GA may introduce bias, with misclassification of preterm and term infants or preterm and small-for-gestational age infants.

Another major measurement issue relates to the timing of exposure. Human reproduction and development by nature is a highly timed and interrelated process (35). It spans across a spectrum of critical windows ranging from pre-pregnancy through to the early neonatal period (36), with any harmful exposures during these critical windows can increase PTB risk. Given this, quantification of the level and timing of exposures is essential for accurate classification of adverse outcomes. Some events are unobservable, so exposure quantification is further complicated by GA imprecision. Inappropriate treatment of exposures in analyses can also introduce bias. Although exposures are typically averaged over the entire pregnancy (by trimester or predefined lag period) in time-fixed analyses, this fails to capture the exposure windows of etiological importance (37). With time-varying exposures such as treatment in pregnancy (38), there is movement in and out of exposure states. Incorrectly treating the exposure as time-fixed ignores actual timing of treatment initiation, and person-time from the study start until delivery contributes wholly to the exposure classification (19, 39). This misclassification of the unexposed time before treatment initiation as exposed, considered as immortal time, can lead to biased estimates (40, 41).

Information Bias: HIV/ART and preterm birth examples

Two studies investigating this association have shown the impact of gestational age measurement error. In South Africa, a birth outcomes study enrolled routine-care women living with and without HIV, with gestational age assessed antenatally using LMP, SFH and ultrasound (42). In a secondary analysis, PTB rates were compared based on assessment methods and their impact on the association with HIV/ART was examined (43). LMP-based GA was found to underestimate gestation relative to ultrasound-based gestational age, contributing to significant differences in preterm birth incidence estimates. An increased risk of PTB by HIV status was observed when GA was assessed using ultrasound, but associations were smaller and not statistically significant when GA assessment was by LMP and SFH. The discrepancy observed in findings between the GA assessment methods was considered to be due to random measurement error. This analysis highlights that the association between maternal ART use and preterm birth may be substantially influenced by GA assessment methods.

In Zambia, a study in routine care enrolled women living with and without HIV, with gestational age assessed using LMP and ultrasound (44); an increased incidence of PTB was observed when GA was assessed using LMP. The study also showed that LMP-based GA estimates were subject to under-

and over-estimation depending on the timing of antenatal care presentation. Error in GA measurement is often thought to be random, as highlighted in the previous study, however these findings showed that LMP-based estimates of preterm among pregnancies may suffer from bias from systematic errors. This suggests that studies that rely on LMP alone are likely to falsely elevate the risk of preterm among groups of women who present later in pregnancy.

Subsequently, the authors of the previously discussed South African and Zambian cohorts conducted a combined analysis showing that ultrasound-based GA estimated PTB incidence was similar in both studies. However, substantial differences were observed in PTB incidence between the two cohorts, when using LMP-based GA: in the Zambian cohort (20.2%) PTB incidence was half of that in the South African cohort (39.7%) (Malaba, personal communication, 2018). In addition to recall issues, factors related to menstrual cycle variations and irregularities also contribute to LMP dating inaccuracies. The authors hypothesised that cohort differences could explain the differences seen with LMP-based estimates, in particular BMI profiles, with a significantly higher proportion of overweight/obese women in the South African cohort than in Zambia cohort (74% vs 41%). Women with higher BMI tend to experience increased menstrual irregularity. This comparative analysis also highlights the need to improve understanding of maternal and fetal factors leading to biased GA estimates.

These examples of studies with unexpected or counterintuitive results highlight the importance of considering and addressing selection and measurement bias when designing and analysing studies investigating the association between PTB and infections and/or their treatments.

3.5 Conclusion

In this article, we briefly reflected on some important methodological considerations that should be regarded when designing studies or information systems for the monitoring or study of PTB, and the interpretation of findings. We described selection and information biases that can arise from inadequate study design, data collection, and analyses procedures, leading to inaccurate findings. The current approach for quantifying the estimates of PTB entails some necessary simplification, because of the limitations related to data availability. As part of achieving SDG 3, reducing PTB and its sequelae in LMICs is essential. We therefore advocate for the strengthening of antenatal care services to improve pregnancy outcomes and the quality of population data, particularly gestational age. The accuracy of this data is contingent on the availability of robust measurements tools such as ultrasound. However in LMICs availability is limited, therefore novel low-cost measurement tools need to be developed for improving gestational age assessment at the individual level. At the population level, leveraging the strengths of novel ML algorithms can strengthen the accuracy of

PTB prediction and diagnosis (5). These efforts need to be coupled with improvements in birth registration systems, with the use of standardised definitions and classification by gestational age, and/or clinical presentation (spontaneous or medically indicated PTB) in collection and reporting of data. Furthermore, continued capacity building of healthcare staff and researchers is critical, for the strengthening of data collection, management and analytic procedures.

We also need to optimise current study designs to study PTB and its risk factors more accurately, by taking into account the important methodological considerations described in this article. In particular, a better understanding and quantification of the error introduced when using commonly used, but error prone, assessment methods, so that PTB rates can be appropriately adjusted when necessary. Additionally, a nuanced understanding of data limitations is required, without which, results of global estimates or etiological research need to be interpreted with caution to avoid incorrect conclusions. New data analysis approaches should be explored as they may provide more efficient ways to use existing data to inform policy and practice.

Preterm birth remains a central public health issue, particularly in LMICs which bear the highest global burden. Despite this, the contribution of these populations to PTB data and research is limited. Further research efforts will require high quality epidemiologic and clinical data from these settings to inform development of context-specific interventions for PTB prevention and management.

3.6 References

1. Chawanpaiboon S, Vogel JP, Moller A-B, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health*. 2019;7(1):e37-e46.
2. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet.* 2012;379(9832):2162-2172.
3. WHO. Global Action Plan: For health lives and well-being for all. Geneva, Switzerland: World Health Organization; 2018.
4. Villar J, Papageorghiou AT, Knight HE, et al. The preterm birth syndrome: a prototype phenotypic classification. *Am J Obstet Gynecol.* 2012;206(2):119-123.
5. Rittenhouse KJ, Vwalika B, Keil A, et al. Improving preterm newborn identification in low-resource settings with machine learning. *PLoS One*. 2019;14(2):e0198919.
6. Okonjo-Iweala N, Osafo-Kwaako P. Improving health statistics in Africa. *Lancet*. 2007;370(9598):1527-1528.
7. Ramakrishnan R, Rao S, He JR. Perinatal health predictors using artificial intelligence: A review. *Womens Health*. 2021;17:17455065211046132.
8. Davidson L, Boland MR. Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes. *Brief Bioinform*. 2021;22(5): bbaa369
9. Simmons LE, Rubens CE, Darmstadt GL, Gravett MG. Preventing preterm birth and neonatal mortality: exploring the epidemiology, causes, and interventions. *Semin Perinatol.* 2010: 34(6):408-415
10. Steer P. The epidemiology of preterm labour. *Br J Obstet Gynaecol*. 2005;112(Suppl 1):1-3.
11. UNAIDS. HIV Prevalence: Adults (15 - 49). Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2021.
12. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737
13. Zash R, Jacobson DL, Diseko M, et al. Comparative safety of antiretroviral treatment regimens in pregnancy. *JAMA Pediatr.* 2017;171(10):e172222
14. Malaba TR, Phillips T, Le Roux S, et al. Antiretroviral therapy use during pregnancy and adverse birth outcomes in South African women. *Int J Epidemiol*. 2017;46(5):1678-1689.
15. Messerlian C, Basso O. Cohort studies in the context of obstetric and gynecologic research: a methodologic overview. *Acta Obstet Gynecol Scand*. 2018;97(4):371-379.

16. Brocklehurst P, French R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *Br J Obstet Gynaecol.* 1998;105(8):836-848.
17. Alexander LK, Lopes B, Ricchetti-Masterson K, Yeatts KB. Selection Bias. ERIC Notebook. 2nd ed. Chapel Hill (NC): The University of North Carolina at Chapel Hill 2015.
18. Snowden JM, Bovbjerg ML, Dissanayake M, Basso O. The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection. *Curr Epidemiol Rep*. 2018;5(4):379–387.
19. Neophytou AM, Kioumourtzoglou M-A, Goin DE, Darwin KC, Casey JA. Addressing special cases of bias that frequently occur in perinatal epidemiology. *Int J Epidemiol*. 2021;50(1):337-345.
20. Moller A-B, Petzold M, Chou D, Say L. Early antenatal care visit: a systematic analysis of regional and global levels and trends of coverage from 1990 to 2013. *Lancet Glob Health*. 2017;5(10):e977-e983.
21. Ananth CV, Schisterman EF. Confounding, causality, and confusion: the role of intermediate variables in interpreting observational studies in obstetrics. *Am J Obstet Gynecol*. 2017;217(2):167-175.
22. Dadabhai S, Gadama L, Chamanga R, et al. Pregnancy outcomes in the era of universal antiretroviral treatment in sub-Saharan Africa (POISE Study). J *Acquir Immune Defic Syndr.* 2019;80(1):7-14.
23. Ramokolo V, Goga AE, Lombard C, Doherty T, Jackson DJ, Engebretsen IM. In utero ART exposure and birth and early growth outcomes among HIV-exposed uninfected infants attending immunization services: results from national PMTCT surveillance, South Africa. *Open Forum Infect Dis*. 2017;4(4):ofx187
24. Slogrove AL, Powis KM, Zash R. Methodological considerations in evaluating pregnancy outcomes in women living with HIV. *J Acquir Immune Defic Syndr.* 2019;81(2):e63-e64
25. Uthman OA, Nachega JB, Anderson J, et al. Timing of initiation of antiretroviral therapy and adverse pregnancy outcomes: a systematic review and meta-analysis. *Lancet HIV.* 2017;4(1):e21-e30.
26. Stoner MCD, Cole SR, Price JT, Winston J, Stringer JSA. Timing of initiation of antiretroviral therapy and risk of preterm birth in studies of HIV-infected pregnant women: The role of selection bias. *Epidemiology*. 2018;29(2):224-229.
27. Chi BH, Vwalika B, Killam WP, et al. Implementation of the Zambia electronic perinatal record system for comprehensive prenatal and delivery care. *Int J Gynaecol Obstet.* 2011;113(2):131-136.
28. Alexander LK, Lopes B, Ricchetti-Masterson K, Yeatts KB. Sources of systematic error or bias: Information bias. ERIC Notebook 2nd ed Chapel Hill (NC): The University of North Carolina at Chapel Hill. 2015.

29. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Berlin, Germany: Springer Science & Business Media; 2011.
30. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol*. 2007;21(Suppl 2):86-96
31. Savitz DA, Terry JW, Dole N, Thorp JM, Siega-Riz AM, Herring AH. Comparison of pregnancy dating by last menstrual period, ultrasound scanning and their combination. *Am J Obstet Gynecol.* 2002;187(6):1660-1666.
32. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol.* 1996;174(1):278-281.
33. Mongelli M, Gardosi J. Symphysis-Fundus Height and Pregnancy Characteristics in Ultrasound-Dated Pregnancies. *Obstet Gynecol.* 1999;94(4):591-594.
34. Alexander GR, de Caunes F, Hulsey TC, Tompkins ME, Allen M. Validity of postnatal assessments of gestational age: a comparison of the method of Ballard et al. and early ultrasonography. *Am J Obstet Gynecol.* 1992;166(3):891-895.
35. Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic measures of the course and outcome of pregnancy. *Epidemiol Rev*. 2002;24(2):91-101.
36. Buck Louis GMP, Platt RW. Reproductive and perinatal epidemiology. Oxford University Press: New York, USA; 2011.
37. Sheridan P, Ilango S, Bruckner TA, Wang Q, Basu R, Benmarhnia T. Ambient Fine Particulate Matter and Preterm Birth in California: Identification of Critical Exposure Windows. *Am J Epidemiol.* 2019;188(9):1608-1615.
38. Corrao G, Rea F, Franchi M, Beccalli B, Locatelli A, Cantarutti A. Warning of Immortal Time Bias When Studying Drug Safety in Pregnancy: Application to Late Use of Antibiotics and Preterm Delivery. *Int J Environ Res Public Health*. 2020;17(18):6465-6476.
39. Pazzagli L, Linder M, Zhang M, et al. Methods for time-varying exposure related problems in pharmacoepidemiology: An overview. *Pharmacoepidemiol Drug Saf*. 2018;27(2):148-160
40. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167(4):492-499.
41. Matok I, Azoulay L, Yin H, Suissa S. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol.* 2014;100(9):658-662.
42. Malaba TR, Myer L, Gray CM, Newell ML. Cohort profile: Prematurity Immunology in Mothers living with HIV and their infants Study (PIMS). *BMJ Open*. 2021;11(9):e047133.
43. Malaba TR, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol*. 2018;28(12):893-900.

44. Price JT, Winston J, Vwalika B, et al. Quantifying bias between reported last menstrual period and ultrasonography estimates of gestational age in Lusaka, Zambia. *Int J Gynaecol Obstet*. 2019;144(1):9-15.

45. Bengtson AM, Chibwesha, CJ, Westreich D, et al. Duration of cART Before Delivery and Low Infant Birthweight Among HIV-Infected Women in Lusaka, Zambia. *J Acquir Immune Defic Syndr*. 2016;71(5):563-569.

46. Boer K, Nellen JF, Patel D, et al. The AmRo study: pregnancy outcome in HIV-1-infected women under effective highly active antiretroviral therapy and a policy of vaginal delivery. *Br J Obstet Gynaecol*. 2007;114(2):148-155.

47. Kesho Bora Study Group, de Vincenzi I. Triple antiretroviral compared with zidovudine and single-dose nevirapine prophylaxis during pregnancy and breastfeeding for prevention of mother-to-child transmission of HIV-1 (Kesho Bora study): a randomised controlled trial. *Lancet Infect Dis*. 2011;11(3):171-180.

48. Njom Nlend AE, Nga Motazé A, Moyo Tetang S, et al. Preterm Birth and Low Birth Weight after In Utero Exposure to Antiretrovirals Initiated during Pregnancy in Yaoundé, Cameroon. *PLoS One*. 2016;11(3):e0150565.

49. Olagbuji BN, Ezeanochie MC, Ande AB, Oboro VO. Obstetric and perinatal outcome in HIV positive women receiving HAART in urban Nigeria. *Arch Gynecol Obstet*. 2010;281(6):991-994.

50. Tuomala RE, Watts DH, Li D. Improved obstetric outcomes and few maternal toxicities are associated with antiretroviral therapy, including highly active antiretroviral therapy during pregnancy. *J Acquir Immune Defic Syndr.* 2005;38(4):449-473.

51. Zash R, Souda S, Chen JY, et al. Reassuring Birth Outcomes with Tenofovir/Emtricitabine/ Efavirenz used for Prevention of Mother to Child Transmission of HIV in Botswana. *J Acquir Immune Defic Syndr.* 2016;71(4):428-436.

52. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.

53. Short CE, Douglas M, Smith JH, Taylor GP. Preterm delivery risk in women initiating antiretroviral therapy to prevent HIV mother-to-child transmission. *HIV Med*. 2014;15(4):233-238.

54. Thorne C, Patel D, and Newell ML. Increased risk of adverse pregnancy outcomes in HIV-infected women treated with highly active antiretroviral therapy in Europe. *AIDS*. 2004;18(7):2337-2339.

55. Townsend CL, Cortina-Borja M, Peckham CS, Tookey PA. Antiretroviral therapy and premature delivery in diagnosed HIV-infected women in the United Kingdom and Ireland. *AIDS.* 2007;21(8):1019-1026.

56. Ekouevi DK, Coffie PA, Becquet R, et al. Antiretroviral therapy in pregnant women with advanced HIV disease and pregnancy outcomes in Abidjan, Cote d'Ivoire. *AIDS*. 2008;22(14):1815–1820.
57. Li N, Sando MM, Spiegelman D, et al. Antiretroviral therapy in relation to birth outcomes among HIV-infected women: a cohort study. *J Infect Dis*. 2016;213(7):1057-1064.
58. Tuomala RE, Shapiro DE, Mofenson LM, et al. Antiretroviral therapy during pregnancy and the risk of an adverse outcome. *N Engl J Med*. 2002;364(24):1863-1870.
59. Slama R, Ballester F, Casas M, et al. Epidemiologic tools to study the influence of environmental factors on fecundity and pregnancy-related outcomes. *Epidemiol Rev*.2014;36(1):148-164.

Chapter 4: Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa.

Malaba TR, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol.* 2018; 28(12):893-900.

Relevance of this paper to the thesis

This chapter presents data on the impact of different gestational age assessment methods on the incidence of preterm delivery and small for gestational age infants. Additionally, there is exploration of the impact of these methods on the association between maternal HIV and antiretroviral therapy use and adverse birth outcomes. The use of different assessment methods has previously been hypothesised as a methodological issue that could explain the heterogeneous findings in previous studies.

Contribution of the student and co-authors

TRM conceptualised and conducted all analyses, led data interpretation and drafted the manuscript. TRM directed data collection, with assistance from HM and AP. MLN and LM conceptualised the study from which these data arise, were responsible for funding, implementation and overall leadership, and assisted with data interpretation. All authors reviewed and approved the final manuscript.

4.1 Abstract

4.1.1 Background

Heterogeneous findings exist on the association between maternal antiretroviral therapy (ART) use and preterm delivery (PTD), or infants born small-for-gestational age (SGA). Whether reported differences may be explained by methods used to ascertain gestational age (GA) has not been explored.

4.1.2 Methods

We enrolled consecutive pregnant women attending a large primary care antenatal clinic in South Africa. Public-sector midwives assessed GA by last menstrual period (LMP) and symphysis-fundal height (SFH). Separately, if clinical GA was <24weeks, ultrasound (US) was performed by a research sonographer blinded to midwife assessments. In analysis, the impact of measurement error on the association between HIV/ART status and birth outcome by GA method was assessed, and factors associated with clinical GA under- or over-estimation identified.

4.1.3 Results

In 1787 women included overall, estimated PTD incidence was 36% by LMP, 17% by SFH and 11% by US. PTD risk was higher for women living with HIV than those without HIV using US-GA (adjusted odds ratio (AOR) 1.95; 95% CI 1.10-3.46); for LMP/SFH-GA the associations were smaller and not significant. These findings persisted after adjustment for age, parity, height and previous PTD. PTD risk did not vary by timing of ART initiation (before or during pregnancy) for any method. Elevated BMI and older age were associated with decreased risk of under-estimation by both LMP and SFH; HIV status and obesity were associated with increased risk of over-estimation by SFH. There were no differences in SGA incidence across GA methods.

4.1.4 Conclusions

Findings for an association between HIV/ART and birth outcomes are substantially influenced by GA assessment method. With growing public health interest in this association, future research efforts should seek to standardise optimal measures of gestation.

4.2 Introduction

Increasing numbers of pregnant women living with HIV use lifelong antiretroviral therapy (ART); in high prevalence HIV countries ART may be the most common drug exposure in pregnancy (1, 2). While ART is invaluable for the health of women living with HIV and their infants, some studies have reported associations between ART use in pregnancy and adverse birth outcomes such as preterm delivery (PTD) and small-for-gestational age (SGA) infants (3-5). PTD and SGA substantially increase the risk of infant mortality, with PTD accounting for the highest proportion of deaths in the critical neonatal period (6). Thus in high HIV prevalence settings it is essential to understand how ART use may influence the risk of preterm delivery. Findings on the association between ART use and adverse birth outcomes, including PTD, are conflicting and possibly associated with ART regimen used (7, 8), timing of initiation and ART eligibility criteria.

Studies investigating ART use in pregnancy are heterogeneous in their gestational age (GA) assessment method. GA assessment is challenging in resource-limited settings where ultrasonography is usually unavailable, and instead routinely based on clinical assessment via dating the last menstrual period (LMP) and/or measurement of the symphysis-fundal height (SFH). Although generally considered most reliable, ultrasound (US) is less accurate for pregnancy dating if carried out after 24 weeks, when biological variability of fetal biometry increases (9, 10). The reliability of LMP-GA is limited by errors related to irregular menstrual cycles of varying duration or accurate recall of LMP dates, and variation in timing of ovulation (11, 12). SFH-GA assessment is difficult early in pregnancy (<12 weeks); and even when measurable its accuracy can be diminished by multiple gestations, high body mass index (BMI), intrauterine growth restriction and other factors (13, 14).

Differences in GA assessment could provide an important methodological explanation for the observed heterogeneity in findings but has not so far been evaluated. Measurement error of GA could lead to biased estimates of the association between HIV/ART and the GA-based outcomes PTD and SGA, with the severity and nature of this bias depending on the form of the error.

We set out to examine the impact of GA assessment method, and any measurement error therein, on the association between HIV/ART status and adverse birth outcomes among a population of pregnant women undergoing routine antenatal care at a large public sector primary care facility in Cape Town, South Africa.

4.3 Methods

Between April 2014 and October 2016, consecutive pregnant women (aged ≥18 years) were enrolled in a prospective cohort (n=3972). All women seeking antenatal care (ANC) at a large public sector primary care facility serving a low-income, high HIV prevalence sub-district of Cape Town South Africa (15). Eligibility was regardless of their HIV status and included women living with HIV who conceived while on ART continued their current regimen, and those newly diagnosed or not on ART were initiated on the WHO recommended first-line regimen. Women clinically assessed (LMP and/or SFH) during routine ANC services to be ≤24 weeks gestation were referred for a same day ultrasound by a research sonographer using standardised assessment protocols and blinded to the midwife GA assessment. The study was reviewed and approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee and the University of Southampton Faculty of Medicine Ethics Committee. Written informed consent was obtained for study participation, including data abstraction from routine clinical records through the pregnancy and post-partum period.

Following their first ANC visit, enrolled women had data abstracted on HIV status, pregnancy history (previous and index pregnancies), medications prescribed during pregnancy and any maternal diagnoses during pregnancy. Obstetric outcomes, including date of delivery and birthweight, were abstracted following delivery. HIV status was categorised as living without HIV or living with HIV; ART status was categorised as initiation before or during pregnancy. The primary outcome was GA at delivery based on completed weeks by LMP, SFH and US. PTD was defined as delivery at <37 weeks' gestation, late to moderate preterm as delivery 32 to 37 weeks and post-term as delivery at >42 weeks (16). Using the gender-specific Intergrowth Standards-21st project standards (17), infants with birthweights <10th percentile for GA were classified as SGA, between 10th and 90th percentile as appropriate-for-gestational age (AGA), and >90th percentile as large-for-gestational age (LGA).

Statistical analyses were performed using STATA version 14.0 (Stata Corporation, College Station, TX, USA). Analyses focused on two comparisons: women living with HIV vs living without HIV; and among women living with HIV ART initiation before pregnancy vs initiation during pregnancy. Outcome variables (PTD and SGA) were created for each assessment method (LMP, SFH and US), using GA at booking and date of delivery. Analyses compared GA estimated by each assessment method in live singleton births. Further analyses were restricted to women with GA estimated by all three assessment methods and to women with both a clinical GA estimate (LMP and/or SFH) and an ultrasound GA estimate. Comparisons of proportions were based on chi-squared tests.

In women with GA by all three assessment methods the associations between HIV/ART status and PTD were examined by each assessment method using logistic regression. Results were presented as odds ratios (OR) with 95% confidence intervals (CI), and adjusted for age, parity, BMI and previous PTD based on their association with adverse birth outcomes (3,5). BMI, assessed at the first ANC visit, was classified as underweight (≤18.5 kg/m^2), normal (18.6-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), or obese (≥30kg/m^2) (18). To inform understanding of GA under- or over-estimation by LMP and SFH compared to US, multinomial logistic regression was conducted. Concordance was defined as <7 days difference which is deemed to be clinically relevant; however for US conducted in the second trimester LMP concordance was defined according to American College of Obstetricians and Gynaecologists recommendations: <7days between 14-15 weeks GA, <10 days between 16-21 weeks and <14 days between 22-27 weeks (19). Results were presented as risk ratios (RR) with 95% CI, and adjusted for age, parity, previous PTD, BMI and ART regimen (in 'living with HIV' only comparisons).

4.4 Results

4.4.1 Demographics

A total of 1787 women with live singleton births were included: 1014 living without HIV (57%) and 773 living with HIV (43%), of whom 368 (48%) initiated ART before pregnancy and 405 (52%) during pregnancy. In line with local and WHO treatment guidelines, all enrolled women living with HIV were on ART, and most (94%) were on a regimen of two NRTIs [tenofovir (TDF) and emtricitabine (FTC)], with the NNRTI efavirenz (EFV). Women living with HIV were older and less likely to be primigravid than women living without HIV (Table 4.1). Women living with HIV who initiated ART before pregnancy were older and less likely to be primigravid than those who initiated during pregnancy. Overall, 28% of women were overweight and 46% obese; there were no significant differences by HIV status or timing of ART initiation.

In total, 88% (n=1585) of women had an LMP-based GA, 68% (n=1220) SFH-based, and 67% (n=1204) US-based, with estimated median GA at enrolment varying by assessment method (Table 4.1). Of the ultrasounds 719 (60%) were conducted in the second trimester with 140 resulting in dating of >24 weeks GA.

Table 4.1 Description of enrolled women (n=1787) at 1st ANC Visit

Maternal characteristics	Total $n = 1787$ (%)	HIV-uninfected $n = 1014$ (%)	HIV-infected $n = 773$ (%)	HIV-infected	
				Initiated before pregnancy $n = 368$ (%)	Initiated during pregnancy $n = 405$ (%)
Age, y					
<24	536 (30)	398 (39)	138 (18)	43 (12)	95 (23)
25–29	500 (28)	291 (29)	209 (27)	76 (21)	133 (33)
>30	751 (42)	325 (32)	426 (55)	249 (67)	177 (44)
Median	28 (24–32)	26 (22–31)	30 (26–33)	32 (28–35)	29 (25–32)
Height, cm					
≤155	500 (28)	270 (27)	230 (30)	108 (29)	122 (30)
156–161	651 (36)	375 (37)	276 (36)	143 (39)	133 (33)
≥162	636 (36)	369 (36)	267 (35)	117 (32)	150 (37)
Median	158 (155–163)	159 (155–163)	158 (154–163)	158 (154–162)	158 (154–163)
Gestation at enrollment, wk					
Determined by LMP	1585	896	689	368	405
Median	17 (12–22)	18 (13–24)	15 (11–21)	14 (11–19)	16 (12–22)
Determined by SFH	1219	709	510	215	295
Median	23 (18–28)	24 (19–28)	20 (16–26)	19 (14–25)	21 (17–27)
Determined by US	1204	603	601	279	322
Median	16 (12–21)	17 (12–22)	15 (11–20)	14 (11–18)	16 (12–20)
Body mass index					
Underweight	16 (1)	10 (1)	6 (1)	5 (1)	1 (0.3)
Normal	329 (18)	188 (19)	141 (18)	66 (18)	75 (19)
Overweight	494 (28)	282 (28)	213 (18)	104 (28)	110 (27)
Obese	823 (46)	467 (46)	356 (46)	166 (45)	189 (47)
Gravidity					
1	383 (21)	263 (26)	120 (16)	41 (11)	79 (20)
2	639 (36)	372 (37)	267 (35)	119 (32)	148 (37)
≥3	765 (43)	379 (37)	386 (50)	208 (57)	178 (44)
Median	2 (2–3)	2 (1–3)	2 (2–3)	2 (2–3)	2 (2–3)
Parity					
0	472 (26)	303 (30)	169 (22)	67 (18)	102 (25)
1	700 (39)	397 (39)	303 (39)	135 (37)	168 (42)
≥2	615 (35)	314 (31)	301 (39)	166 (45)	135 (33)
Median	1 (0–2)	1 (0–2)	1 (0–2)	1 (1–2)	1 (1–2)
Previous preterm*					
Yes	127 (9)	58 (8)	69 (11)	37 (11)	32 (10)
Hemoglobin g/dL					
Normal (≥11.0)	792 (44)	496 (49)	296 (46)	171 (46)	125 (31)
Mild Anemia (9–10.9)	613 (34)	337 (33)	276 (43)	111 (30)	165 (41)
Moderate Anemia (7–8.9)	122 (7)	54 (5)	68 (11)	25 (7)	43 (11)
Severe Anemia (<7)	3 (0.2)	3 (0.3)	0	0	0
ART regimen					
TDF-3TC-EFV	691 (39)	—	691 (94)	298 (89)	393 (98)
TDF-3TC-NVP	4 (0.2)	—	4 (1)	2 (1)	2 (0.5)
Other NNRTI	20 (1)	—	20 (2)	14 (4)	6 (1)
PI regimen	22 (1)	—	22 (3)	20 (6)	2 (0.5)

*among women with a previous pregnancy

All variables, with the exception of height, BMI and haemoglobin had <1% missing data. For height, 6% (n=111) of data was missing and for BMI, 7% (n=126) was missing, for both there were similar proportions across all comparison groups. For Haemoglobin, 14% (n=57) of data was missing with similar proportions across all comparison groups. For ART status 2% (n=36) were missing regimen data

4.4.2 Adverse birth outcomes

4.4.2.1 Incidences of adverse birth outcomes

Among all women

The estimated incidence of PTD using LMP-GA was 36% (95% CI 33-39%), 17% (95% CI 15-19%) by SFH and 11% (95% CI 9-13%) by US; with no significant difference by HIV status using LMP-GA (37% vs 36%) or SFH-GA (18% vs 17%). However, US-GA estimated PTD incidence was significantly higher in women living with HIV than those living without HIV (14% vs 8%) (Table 4.2; Figure 4.1a). No significant differences in PTD incidence were observed between women living with HIV initiating ART before or during pregnancy by LMP-GA (35% vs 39%), SFH-GA (18% vs 17%) and US-GA (12% vs 15%) (Table 4.2; Figure 4.1a).

Table 4.2 Birth outcomes in overall cohort (n=1787)

Birth outcome	Total	HIV-uninfected	HIV-infected	HIV-infected	
				Initiated before pregnancy	Initiated during pregnancy
GA (wks) LMP	$n = 1585$	$n = 896$	$n = 689$	$n = 336$	$n = 353$
Post-term >42	115 (7)	57 (6)	58 (8)	33 (10)	25 (7)
Term (37–42)	896 (57)	522 (58)	374 (55)	185 (55)	189 (54)
Any preterm (<37)*	574 (36)	317 (36)	257 (37)	118 (35)	139 (39)
Late to mod preterm (32–37)	418 (73)	238 (75)	180 (70)	86 (73)	94 (68)
Very preterm (28–32)	156 (27)	79 (25)	77 (30)	32 (27)	45 (32)
GA (wks) SFH	$n = 1219$	$n = 709$	$n = 510$	$n = 215$	$n = 295$
Post-term >42	97 (8)	59 (8)	38 (7)	15 (7)	23 (8)
Term (37–42)	914 (75)	532 (75)	382 (75)	161 (75)	221 (75)
Any preterm (<37)*	208 (17)	118 (17)	90 (18)	39 (18)	51 (17)
Late to mod preterm (32–37)	171 (82)	98 (83)	73 (81)	33 (85)	40 (78)
Very preterm (28–32)	37 (18)	20 (17)	17 (19)	6 (15)	11 (22)
GA (wks) US	$n = 1204$	$n = 603$	$n = 601$	$n = 279$	$n = 322$
Post-term >42	6 (1)	6 (10)	0	0	0
Term (37–42)	1065 (88)	547 (91)	518 (86)	244 (88)	274 (85)
Any Preterm (<37)*	133 (11)	50 (8)	83 (14)	35 (12)	48 (15)
Late to mod preterm (32–37)	114 (86)	47 (94)	67 (81)	28 (80)	39 (81)
Very preterm (28–32)	19 (14)	3 (6)	16 (19)	7 (20)	9 (19)
Size for GA LMP	$n = 1441$	$n = 825$	$n = 616$	$n = 294$	$n = 322$
Small (SGA)	137 (10)	61 (7)	76 (12)	36 (12)	40 (13)
Appropriate (AGA)	847 (59)	498 (60)	349 (57)	181 (62)	168 (52)
Large (LGA)	457 (32)	266 (32)	191 (31)	77 (26)	114 (35)
Size for GA SFH	$n = 1113$	$n = 649$	$n = 464$	$n = 195$	$n = 269$
Small (SGA)	145 (13)	72 (11)	73 (16)	22 (11)	51 (19)
Appropriate (AGA)	814 (73)	481 (74)	333 (72)	150 (77)	183 (68)
Large (LGA)	154 (14)	96 (15)	58 (13)	23 (12)	35 (13)
Size for GA US	$n = 1184$	$n = 596$	$n = 588$	$n = 273$	$n = 315$
Small (SGA)	131 (11)	49 (8)	82 (14)	39 (14)	43 (14)
Appropriate (AGA)	944 (80)	492 (83)	452 (77)	213 (78)	239 (76)
Large (LGA)	109 (9)	55 (9)	54 (9)	21 (8)	33 (10)

* Any Preterm consists of the subset of Late Preterm, Moderately Preterm and Very Preterm deliveries

For Size for GA data missing for women whose GA is not within Intergrowth standard limits (<24w and >42w) and those missing either GA or birthweight

Among women with GA assessed by all three methods (Restricted Group)

Among the 35% (n=629) of women with GA assessed by all three methods, estimated median (IQR) GA at enrolment was 17 weeks (14-21) by LMP, 20 weeks (16-23) by SFH and 19 weeks (15-23) by US (Supp. Table 9.2.1a). HIV status was the only variable associated with having been assessed by all three methods, with increased odds for women living with HIV (OR 1.54; 95% CI 1.22 - 1.84). In this group, estimated PTD incidence was 42% (95% CI 38-46%) using LMP-GA, 14% (95% CI 12-17%) using SFH-GA and 10% (95% CI 7-12%) using US-GA (Supp. Table 9.2.2). As before for the whole group, estimated PTD incidence was significantly higher in women living with HIV than in those living without HIV using US-GA (12% vs 7%) but not significantly different using LMP-GA (40% vs 44%) or SFH-GA (16% vs 12%) (Supp. Table 9.2.2; Figure 4.1b). In women living with HIV (n=320), no significant differences were observed between women initiating ART before or during pregnancy for LMP-GA (39% vs 41%), SFH-GA (17% vs 15%) or US-GA (10% vs 15%) (Supp. Table 9.2.2; Figure 4.1b).

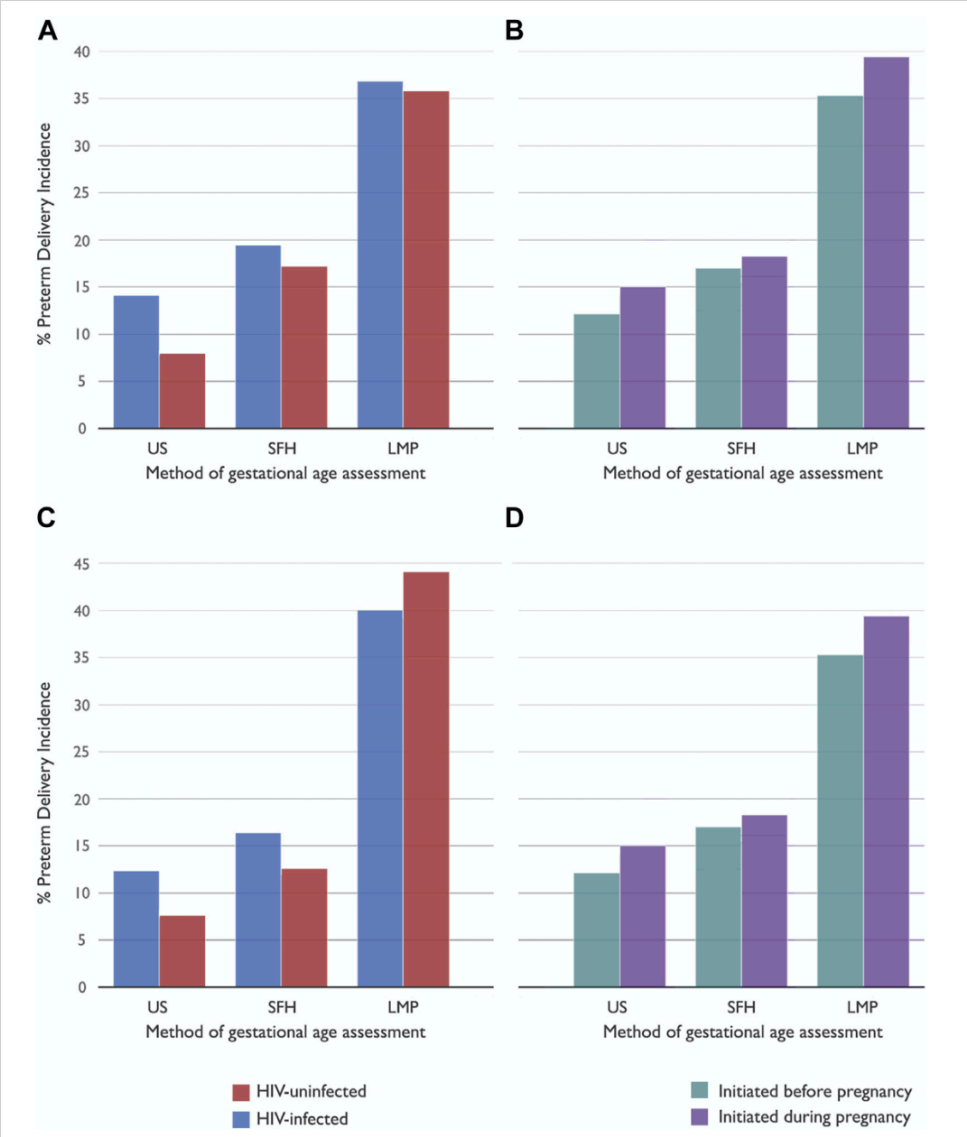

Figure 4.1 Incidence of preterm deliveries (<37weeks) by HIV/ART Status

A in all enrolled women (n=1787), and **B** in all women living with HIV (n=773), **C** in women assessed by all three methods (n=629), and **D** in women living with HIV assessed by all three methods (n=320)

Adjusting for age, parity and BMI, women living with HIV had increased odds of US-GA PTD (AOR 1.98, 95% CI 1.12-3.52) compared to women living without HIV. This was also seen with SFH-GA although not statistically significant (AOR 1.34, 95% CI 0.83-2.15); associations seen with LMP-GA were small and not significant (Table 4.3). In women living with HIV, the odds of PTD did not vary appreciably by timing of ART initiation across the three assessment methods (Table 4.3). Similar results were seen in the overall cohort (Supp. Table 9.2.3).

Table 4.3 Adjusted association between HIV/ART status and preterm delivery in women

Assessment method	HIV-infected versus HIV-uninfected (Ref)*			Initiated before pregnancy (Ref) versus initiated during pregnancy†		
		aOR (95% CI)	*P*-value		aOR (95% CI)	*P*-value
LMP	HIV-infected	0.89 (0.63–1.24)	.483	During pregnancy	0.94 (0.56–1.60)	.802
SFH	HIV-infected	1.34 (0.83–2.15)	.225	During pregnancy	1.42 (0.76–2.84)	.290
US	HIV-infected	1.98 (1.12–3.53)	.020	During pregnancy	0.69 (0.33–1.47)	.335

*Adjusted for age, parity, BMI and previous PTD
**Adjusted for age, parity, BMI, previous PTD and ART regimen[4]

4.4.2.2 Factors associated with under- and over-estimation of GA

In women with both an US estimate and a clinical estimate LMP (n=1057) or SFH (n=732), the estimated concordance with US-GA estimates was 29% (n=308) for LMP and 36% (n=261) for SFH. LMP under-estimated US-GA in 484 cases (46%), and over-estimated in 265 cases (25%). The percentages of under- and over-estimation for SFH were 21% (n=153) and 43% (n=318) respectively.

Comparing LMP-GA to US-GA, HIV status was not significantly associated with under-estimating (risk ratio (RR) 0.96, 95% CI 0.71-1.27), but was associated with over-estimating (RR 1.33, 95% CI 0.95-1.64) GA, although not statistically significant. These associations persisted in adjusted models (Table 4.4). No other factors were associated with under- or over-estimating by LMP-GA.
In women living with HIV, timing of ART initiation was not significantly associated with either under- (RR 0.87, 95% CI 0.53-1.45) or over-estimating (RR 1.34, 95% CI 0.78-2.31) GA by LMP and this persisted in the adjusted models (Table 4.4).

Comparing SFH-GA to US-GA, there was no association with HIV status for either under- (RR 1.21, 95% CI 0.80-1.83) or over-estimating (RR 1.33, 95% CI 0.94-1.87) GA. In the adjusted model, obese women (aRR 1.72, 95% CI 1.09-2.73) were at increased risk of over-estimation by SFH; while being older (>30yrs) was associated with decreased risk of under-estimation by SFH (aRR 0.58, 95% CI 0.32-1.03). Among women living with HIV there was no association with timing of ART initiation for either under- or over-estimating GA by SFH and this persisted with adjustment (Table 4.4).

[4] Analysis re-run excluding previous preterm delivery as an adjustment variable, new table included in supplementary material of thesis (Suppl Table 9.2.5).

Table 4.4 Adjusted association between HIV/ART status and concordance with US-based measures in women with both an US estimate and a clinical estimate

Assessment method	HIV-infected versus HIV-uninfected (Ref)*				Initiated before pregnancy (Ref) versus initiated during pregnancy†			
	Ref category: No difference		aRR (95% CI)	*P*-value	Ref category: No difference		aRR (95% CI)	*P*-value
LMP ($n = 1057$)	HIV-infected	Underestimate	0.96 (0.71–1.27)	.764	During pregnancy	Underestimate	0.87 (0.53–1.45)	.609
		Overestimate	1.33 (0.95–1.64)	.098		Overestimate	1.34 (0.78–2.31)	.288
SFH ($n = 732$)	HIV-infected	Underestimate	1.21 (0.80–1.83)	.364	During pregnancy	Underestimate	1.23 (0.66–2.30)	.518
		Overestimate	1.33 (0.94–1.87)	.102		Overestimate	0.91 (0.54–1.52)	.711

*Adjusted for age, parity, BMI
**Adjusted for age, parity, BMI and ART regimen

Figure 4.2 Birthweight and gestational age according to GA assessment method (boys and girls separately)

4.4.2.3 Small for gestational age deliveries

The incidence of SGA was similar according to LMP, SFH and US (10% vs 13% vs 11%); however differences were observed by assessment method for AGA and LGA. . Much wider GA estimate variations were observed when using LMP-GA and SFH-GA than using US-GA, leading to more in the distribution of the tails which in turn would result in more AGA and LGA infants (Figure 2). The incidence of AGA was significantly higher with SFH-GA (73%) and US-GA (80%) than with LMP-GA (59%). Correspondingly, the incidence of LGA was significantly higher when LMP-GA was used (32%) than when SFH-GA (14%) and US-GA (9%) was used (Table 4.2; Figure 4.2).

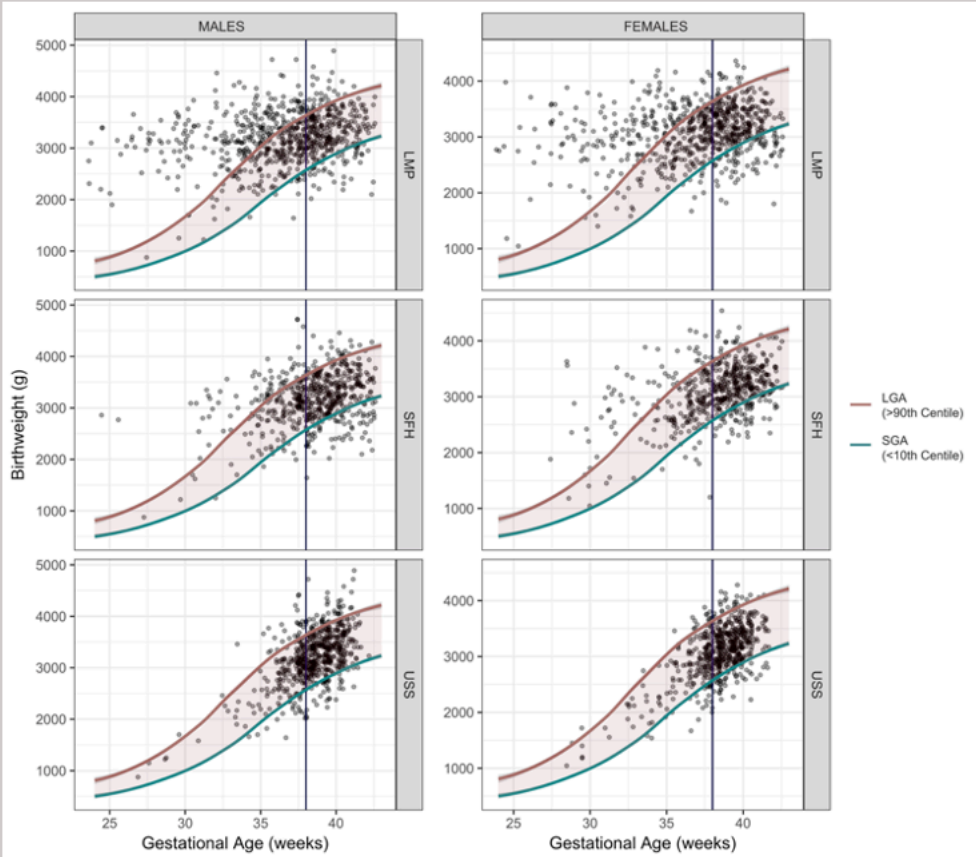

Figure 4.2 Birthweight and gestational age according to GA assessment method (boys and girls separately)

* Superimposed smoothed 10th (SGA) and 90th (LGA) centile curves for birthweight according to gestational age and gender
SGA – small for gestational age; LGA – large for gestational age

When the incidence of SGA was compared according to HIV status, SGA incidence was higher for women living with and without HIV across all three assessment methods. Using LMP-GA or US-GA, SGA estimated did not differ by timing of ART initiation. However, using SFH-GA, women initiating before pregnancy appeared to have a lower SGA incidence than those initiating during pregnancy (11% vs 19%) (Table 4.2). In women with GA

by all three assessment methods, overall SGA was slightly higher when SFH-GA (14%) was used compared to LMP-GA (9%) and US-GA (11%). Similar patterns to those seen overall were seen in this restricted group with the incidence of AGA and LGA.

By all three assessment measures, SGA incidence was higher in women living with and without HIV. No differences were observed according to timing of ART initiation for LMP-GA and US-GA, however a lower incidence of SGA was observed in women initiating before pregnancy for SFH-GA (Supp. Table 9.2.2).

4.5 Discussion

In this prospective cohort of pregnant women seeking ANC at a large public sector primary care facility in South Africa, we found that estimated GA varied by the measurement employed and that differences in GA resulted in significant differences in the occurrence of PTD. Where analysis of safety of ART use in pregnancy by timing of ART initiation is focused on outcomes such as PTD or SGA, our work highlights the need for standardised GA ascertainment methods in these analyses.

Using US-GA, PTD risk was associated with maternal HIV infection, with women living with HIV, all on ART, almost twice as likely to deliver preterm than women living with HIV. A weaker association was observed with SFH-based GA; and minimal differences with LMP-GA. We did not find any appreciable differences in PTD risk for women living with HIV by timing of ART initiation by all three assessment methods. Our findings (in both the overall cohort and in women with all three assessments) suggest that GA ascertainment methods could partially explain the heterogeneity of findings from previous studies on the association between ART use and adverse birth outcomes, suggesting that care should be taken when comparing results from different studies that use different GA measurement methods.

Our finding for a higher PTD incidence in women living with HIV than those living without HIV when using US-GA is consistent with several studies, using different assessment methods, across different settings with the incidence of any PTD (<37 weeks) ranging from 13% to 37% (3, 7, 8, 20, 21). Using US-GA we found an incidence of 11% (95% CI 9-13%) among women living with HIV overall, similar to findings from a Nigerian study (21); but significantly lower than findings from other studies (3, 7, 8, 20, 21). In addition to differences in assessment methods, this could also be due to underlying differences in maternal (HIV disease stage and clinical characteristics) and behavioural risk factors (alcohol and substance abuse).

Using LMP-GA and SFH-GA, we found estimated PTD incidence to be higher than with US-GA, and similar to findings from previous studies, which primarily used LMP-based measures. However, in

adjusted analyses the difference in PTD risk by HIV status was marginal and not significant using either method. The most plausible explanation for the discrepancy in findings between US and LMP/SFH measures is random measurement error. When stratified by HIV and BMI status, the proportions of PTD observed by each assessment method remained similar to the overall, pointing to non-differential misclassification of PTD status due to random measurement error in the GA estimated by LMP and SFH. In previous studies, GA measurement error resulted in an over-estimation of preterm and post-term delivery incidence (9, 12, 22), consistent with our findings of over-estimation by both LMP-GA and SFH-GA compared to US-GA. Our findings show that PTD over-estimation by LMP/SFH was driven primarily by the inflation of the tails of the GA distribution, where estimates were more extreme than the likely true-data estimates.

In our setting, LMP-GA estimates varied widely with measurement error resulting predominantly in under-estimated GA contributing to the substantially higher PTD incidence, in contrast over-estimation would have led to higher frequencies of post-term delivery classification. While we expected some differences between LMP-GA and US-GA (23), the degree of discrepancy was greater than previous reports from high-income countries and other resource-limited settings (23, 24); it was, however, in keeping with previous South African data (25, 26). These inconsistencies suggest the degree and type of error in LMP-GA may depend on demographic characteristics of the study population and data collection methods. For example, errors in GA estimates due to delayed ovulation usually leads to GA over-estimation, whereas GA errors due to recall issues can lead to errors in either direction. Additionally, women with non-normal BMI (<19 or >29kg/m^2) may misreport their LMP due to menstrual irregularities (27, 28). In our study population, numerous factors could have played a part in the under- and over-estimation of LMP-based PTD frequencies (relative to US). In addition to BMI, other socio-demographic factors, known to be associated with reduced accuracy of LMP-GA, are prevalent in our setting. This includes high level of injectable contraception use (29), long-term use of which affects return to menses and fertility, limiting LMP reliability (30). Additionally younger, primigravid women and women with lower educational levels are more likely to misreport LMP (12, 27). Since women living without HIV were more likely to be younger and primigravid it is possible they misreported their LMP more than women living with HIV, which could explain why the expected association by HIV status was not observed with LMP-GA.
In contrast to LMP-GA, in our study SFH-GA was predominantly over-estimated, possibly a reflection of study participants BMI profile. Similar to a previous South African study, SFH-GA performed better than LMP-GA, with accuracy worsening with increasing BMI presumably due to increased maternal adiposity interfering with SFH measurement (25). This is consistent with reports that high BMI affects the accuracy of SFH-GA estimates (13, 31); a Mozambican study demonstrated that among women of the same GA, the fundal height of obese women was higher than that of normal-

weight women resulting in GA over-estimation (31). Accuracy in overweight and obese women is compromised because GA calculation is based on fundal height growth curves derived from normal weight women (32). SFH measurements are also subject to intra- and inter-observer errors, a particular problem with single measurements (32). Multiple SFH measurement models have been shown to have a higher level of accuracy when repeated measurements are used (33). While the association between HIV status and PTD was in the expected direction, attenuation by BMI-related measurement errors could have resulted in non-significant associations.

In women living with HIV, we did not observe any significant association between timing of ART initiation and PTD by any of the assessment methods. An explanation for this is that our sample size was not sufficiently large enough to enable detection of small to moderate differences between these groups of women, however our results are consistent with findings from a previous South African cohort (20), and findings from other settings (34). This suggests that in our study population there might be minimal differences across the comparison group and that any GA measurement errors were equally distributed.

Overall, SGA incidence did not differ greatly by assessment method, which was surprising given GA differences observed according to assessment method. In contrast, LGA estimates differed by assessment methods, with larger than expected proportions of LGA by LMP-GA and SFH-GA, which would misclassification due to inaccurate GA. These results suggest that in our setting, in the absence of ultrasound facilities, SFH-GA is substantially more reliable than LMP-GA.

A major strength of our study was the prospective data collection and GA assessment by a research sonographer, who was blinded to other measures, ensuring a high-quality measure of gestation. Further, US was conducted in women ≤24 weeks, when US is highly reproducible and accurate to date pregnancy duration (35). While there is a possibility that our US-GA estimates may be biased due to systematic under-estimation of GA for smaller fetuses, these errors have been shown to be relatively small compared to the larger errors seen when using LMP- and SFH-based measures for GA estimation (12). While we did not directly measure LMP and/or SFH as part of study procedures and relied instead on data abstraction from routine records, we believe any measurement errors were random rather than systematic. Study limitations include the absence of data on pre-pregnancy BMI and other factors that can influence GA estimate accuracy, such as maternal education which is known to impact LMP-GA estimate accuracy. In addition to measurement error, selection bias is also likely to have contributed to the attenuation of associations in the group of women with three measures. Moreover, given that clinical estimates, which we showed are subject to substantial measurement error, were used to determine which women were referred for ultrasounds there is a

possibility of selection bias which could have led to inaccurate estimate of the true association. Additionally, while it was intended that the sonographer be blinded, this was only partial since by study protocol only women with ≤24 weeks GA were referred for ultrasound assessment.

4.6 Conclusion

The relationship between ART and adverse birth outcomes is complex: it is well established that ART can improve birth and child health outcomes by improving maternal health, but ART may also contribute to worsening outcomes through mechanisms that have not been fully elucidated. With widespread use of ART in pregnancy and growing public health interest in this association, understanding and addressing adverse birth outcomes is vitally important. This study makes an important contribution to this emerging science by demonstrating that inaccurate GA measurement can distort associations between maternal characteristics, including HIV/ART, and adverse birth outcomes. In settings where LMP and SFH-based measures are widely used, bias due to GA ascertainment methods employed should be considered as an alternative explanation for observed associations or null findings. Additionally, given that there appear to be systematic differences in GA estimates according to BMI status, future research efforts should consider the BMI profile of study populations and seek to standardise optimal measures of GA. While early US should be encouraged whenever possible, simple and novel methods that accurately measure GA are required particularly in settings where ultrasound is unavailable; and this should include postnatal assessments to confirm GA based on antenatal determination. Furthermore, quantitative bias analytic techniques that account for measurement error in GA and/or missing data should be considered in future analyses.

4.7 References

1. WHO. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. Geneva, Switzerland: World Health Organization; 2015.
2. UNAIDS. On the fast-track to an AIDS-free generation. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2016.
3. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.
4. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.
5. Li N, Sando MM, Spiegelman D, et al. Antiretroviral therapy in relation to birth outcomes among HIV-infected women: a cohort study. *J Infect Dis*. 2016;213(7):1057-1064.
6. UNICEF. Levels and Trends in Child Mortality. New York, USA: United Nations Childrens Fund; 2015.
7. Boer K, Nellen JF, Patel D, et al. The AmRo study: pregnancy outcome in HIV-1-infected women under effective highly active antiretroviral therapy and a policy of vaginal delivery. *Br J Obstet Gynaecol.* 2007;114(2):148-155.
8. Zash R, Souda S, Chen JY, et al. Reassuring Birth Outcomes with Tenofovir/Emtricitabine/ Efavirenz used for Prevention of Mother to Child Transmission of HIV in Botswana. *J Acquir Immune Defic Syndr.* 2016;71(4):428-436.
9. Gardosi J, Geirsson RT. Routine ultrasound is the method of choice for dating pregnancy. *Br J Obstet Gynaecol.* 1998;105(9):933-936.
10. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol.* 1996;174(1):278-281.
11. Johnson SR, Miro F, Barrett S, Ellis JE. Levels of urinary human chorionic gonadotrophin (hCG) following conception and variability of menstrual cycle length in a cohort of women attempting to conceive. *Curr Med Res Opin*. 2009;25(3):741–748.
12. Dietz PM, England LJ, Callaghan WM, Pearl M, Wier ML, Kharrazi M. A comparison of LMP-based and ultrasound-based estimates of gestational age using linked California livebirth and prenatal screening records. *Paediatr Perinat Epidemiol*. 2007;21(Suppl 2):62-71.
13. Mongelli M, Gardosi J. Sympysis-fundus height and pregnancy characteristic in ultrasound-dated pregnancies. *Obstet Gynecol.* 1999;94(4):591-594.
14. Gardosi J, Francis A. Controlled trial of fundal height measurement plotted on customized antenatal growth charts. *Br J Obstet Gynaecol*. 1999;106(4):309-317.
15. Myer L, Phillips TK, Manuelli V, McIntyre J, Bekker LG, Abrams EJ. Evolution of antiretroviral therapy services for HIV-infected pregnant women in Cape Town, South Africa. *J Acquir Immune Defic Syndr.* 2015;69(2):e57–e65.

16. WHO. Fact Sheet: Preterm Birth. WHO Media Centre 2017.
17. Villar J, Ismail LC, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet.* 2014;384(9946):857-868.
18. WHO. Fact Sheet: HIV/AIDS. Geneva, Switzerland: World Health Organization Media Centre; 2016.
19. American College of Obstetricians and Gynecologists. Methods for estimating the due date. Committee Opinion No. 700. *Obstet Gynecol.* 2017;129(5):e150–e154.
20. Malaba TR, Phillips TK, Le Roux S, et al. Antiretroviral therapy use during pregnancy and adverse birth outcomes in South African women. *Int J Epidemiol*. 2017;46(5):1678-1689
21. Olagbuji BN, Ezeanochie MC, Ande AB, Oboro VO. Obstetric and perinatal outcome in HIV positive women receiving HAART in urban Nigeria. *Arch Gynecol Obstet*. 2010;281(6):991-994.
22. Haglund B. Birthweight distributions by gestational age: comparison of LMP-based and ultrasound-based estimates of gestational age using data from the Swedish Birth Registry. *Paediatr Perinat Epidemiol*. 2007;21(s2):72-78.
23. Tunón K, Eik-Nes SH, Grøttum P. A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15 000 examinations. *Ultrasound Obstet Gynecol*. 1996;8(3):178-185.
24. Jehan I, Zaidi S, Rizvi S, et al. Dating gestational age by last menstrual period, symphysis-fundal height, and ultrasound in urban Pakistan. *Int J Gynaecol Obstet*. 2010;110(3):231-234.
25. Geerts LT, Poggenpoel E, Theron GB. A comparison of pregnancy dating methods commonly used in South Africa: a prospective study. *S Afr Med J*. 2013;103(8):352-357.
26. Geerts LT, Theron AM, Grove D, Theron GB, Odendaal HJ. A community-based obstetric ultrasound service. *Int J Gynaecol Obstet*. 2004;8(1):23-31.
27. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol*. 2007;21(s2):86-96.
28. Kullinger M, Wesström J, Kieler H, Skalkidou A. Maternal and fetal characteristics affect discrepancies between pregnancy-dating methods: a population-based cross-sectional register study. *Acta Obstet Gynecol Scand*. 2017;96(1):86-95.
29. Chersich MF, Wabiri N, Risher K, et al. Contraception coverage and methods used among women in South Africa: A national household survey. *S Afr Med J*. 2017;107(4):307-314.
30. Waller DK, Spears WD, Gu Y, Cunningham GC. Assessing number-specific error in the recall of onset of last menstrual period. *Paediatr Perinat Epidemiol*. 2000;14(3):263–267.
31. Challis K, Osman NB, Nyström L, Nordahl G, Bergström S. Symphysis-fundal height growth chart of an obstetric cohort of 817 Mozambican women with ultrasound-dated singleton pregnancies. *Trop Med Int Health*. 2002;7(8):678-684.

32. Deeluea J, Sirichotiyakul S, Weerakiet S, Arora R, Patumanond J. Fundal height growth curve for underweight and overweight and obese pregnant women in Thai population. *ISRN Obstet Gynecol*. 2013;2013:1-8

33. White LJ, Lee SJ, Stepniewska K, et al. Estimation of gestational age from fundal height: a solution for resource-poor settings. *J R Soc Interface*. 2012;9(68):503-510.

34. Njom Nlend AE, Nga Motazé A, Moyo Tetang S.et al. Preterm Birth and Low Birth Weight after In Utero Exposure to Antiretrovirals Initiated during Pregnancy in Yaoundé, Cameroon. *PLoS One*. 2016;11(3):e0150565.

35. Butt K, Lim K, Diagnostic Imaging Committee. Determination of gestational age by ultrasound. *J Obstet Gynaecol Can*. 2014;36(2):171-181.

Chapter 5: Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age

Malaba TR, Mukonda E, Matjila M, Madlala HP, Myer L, Newell ML, PIMS Study Group. Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age. *Paediatr Perinat Epidemiol*. 2022; 36(4):525-535.

Relevance of this paper to the thesis

This chapter builds on the findings of the analyses in chapter 4, and reports on the association between maternal ART use, timing of ART initiation and adverse birth outcomes. Quantifying adverse birth outcome related to ART exposure in fetal life is essential because an increasing number of women are receiving lifelong antiretroviral therapy, particularly in sub-Saharan Africa, where HIV infection is most common. This chapter ensures that the estimates of preterm delivery, small for gestational age and low birth weight were the most accurate possible, with gestational age based on last menstrual period corrected using multiple overimputation to reduce measurement bias and outcome misclassification.

Contribution of the student and co-authors

TRM conceptualised and conducted all analyses, led data interpretation and drafted the manuscript. TRM also directed data collection, with assistance from HM. EM assisted with part of the data analysis related to multiple overimputation. MLN and LM conceptualised the study from which these data arise, were responsible for funding, implementation and overall leadership, and assisted with data interpretation. All authors reviewed and approved the final manuscript.

5.1 Abstract

5.1.1 Background

Studies of antiretroviral therapy (ART) use during pregnancy have suggested that ART exposure may be associated with a range of adverse outcomes. Efforts to reliably quantify the association between HIV/ART status and adverse birth outcomes have been hampered by inconsistent findings linked to gestational age (GA) assessment bias related issues.

5.1.2 Methods

Pregnant women, HIV-negative or living with HIV (WLHIV), were enrolled at first antenatal care visit and followed through delivery. Ultrasound assessed gestational age was deemed the gold standard. Based on quantitative bias analysis for outcome misclassification, gestational age by non-ultrasound assessment was corrected using multiple overimputation, which deals with missing data and measurement error simultaneously. Using bias-corrected gestational age, birth outcomes were compared between WLHIV and HIV-negative women, and among WLHIV who initiated ART before versus during pregnancy, further divided into trimesters.

5.1.3 Results

Of 3952 women enrolled, 37% were WLHIV (mostly using tenofovir + emtricitabine + efavirenz). Last menstrual period (LMP)-based gestational age was identified to be biased and LMP measures were thus corrected using multiple overimputation. Comparing WLHIV and HIV-negative women, adjusted risk ratio (ARR) of overall pregnancy loss was 1.26 (95% Confidence Interval (CI) 0.98, 1.61); ARR of preterm delivery (PTD) was 1.02 (95% CI 0.88, 1.20); ARR of small for gestational age infants was 1.43 (95% CI 1.14, 1.80). Among WLHIV, outcomes were similar by ART initiation timing.

5.1.4 Conclusions

In this routine care cohort, risk of SGA, and possibly of pregnancy loss, was increased in WLHIV compared to HIV-negative women, with no evidence of increased risk of PTD. Further research is needed to improve mechanistic understanding of the contribution of ART to adverse birth outcomes to optimise treatment for pregnant WLHIV and ensure optimal maternal and infant outcomes.

5.2 Introduction

Sustainable Development Goal 3 target 3.2 aims to reduce deaths of newborns and children under 5 years (1) and continued focus on preterm delivery (PTD), low birthweight (LBW), and small for gestational age (SGA) is warranted. PTD is the second largest cause of under-5 mortality and the most important direct cause of neonatal death, while LBW and SGA are indirect causes of neonatal death (2).

Increasing numbers of women living with HIV (WLHIV), pregnant or not, receive lifelong antiretroviral therapy (ART)(3). HIV status with or without ART use from before or during pregnancy may be associated with adverse birth outcomes. Consequently, quantifying adverse birth outcome possibly relating to ART exposure in fetal life remains vital, particularly in sub-Saharan Africa, where HIV infection is common (4). Increased risk of PTD (5-8), LBW (7-9), and/or SGA (5, 10) infants for WLHIV were reported from some, but not all, studies in sub-Saharan Africa. Reported differences have been posited to arise from ART regimen complexity, drug classes and timing of ART initiation. However, birth outcomes were mostly not the primary study aim, which may have affected data quality, particularly pertinent for gestational age assessment.

First trimester ultrasound is considered the gold-standard for gestational age assessment but is not routinely available in resource-limited settings. We previously showed last menstrual period (LMP) to be an unreliable gestational age measure, with any association between HIV status and ART use with birth outcomes substantially influenced by assessment method (11). Recently, the value of using multiple overimputation to address LMP-based gestational age measurement error was highlighted in a study in Zambia (12). Using this correction method, we aimed to more accurately assess PTD and SGA prevalence and estimate the associations between HIV status, timing of ART initiation and adverse birth outcomes in pregnant women undergoing routine antenatal care at a large public sector primary care facility in Cape Town, South Africa.

5.3 Methods

We conducted an observational cohort study among pregnant women in Cape Town, South Africa from April 2015-October 2016 (PIMS) (13). Briefly, women aged ≥18 years were enrolled consecutively and followed from their 1st antenatal care (ANC) visit to delivery through review and abstraction of routine clinical antenatal and obstetric records.

5.3.1 HIV and treatment

At the 1st ANC visit, women without previous HIV diagnosis were routinely tested. WLHIV who conceived while on ART continued their current regimen; those newly diagnosed and those WLHIV

not on ART were initiated on a fixed dose combination of tenofovir (TDF), emtricitabine (FTC) and efavirenz (EFV). Women testing negative were retested prior to or at delivery; all WLHIV enrolled in PIMS received ART from before pregnancy or from initiation during pregnancy.

5.3.2 Gestational age assessment

At the 1st ANC visit, public sector midwives estimated gestational age for all women using self-reported LMP and/or measurement of the symphysis-fundal height (SFH). LMP gestational age was not assessed in women with uncertain dates, while those without SFH gestational age were assessed at <12 weeks when palpation is difficult. Following study procedures, women clinically assessed to be ≤24 weeks were referred for a same day ultrasound (GE *LOGIQ C3 Premium)* by a research sonographer to determine an ultrasound-based gestation. Using standardised assessment protocols and blinded to the gestational age estimated by the midwife, gestational age from the ultrasound machine estimation was abstracted.

5.3.3 Cohort selection

We included all pregnant WLHIV and HIV-negative women (Figure 5.1). Women with non-singleton pregnancies (n=58) and missing outcomes (n=188) were excluded.

5.3.4 Exposure

The exposure of interest was HIV-infection and timing of ART initiation. HIV status was categorised as HIV-negative or living with HIV; ART status as initiation preconception or during pregnancy, further divided into first (≤13 weeks), second (14-26 weeks) or third (≥27 weeks) trimester. ART regimens were categorised as TDF+FTC+EFV, other non-nucleoside reverse transcriptase inhibitors (NNRTI) regimens (TDF+FTC+NVP or regimens containing other NNRTIs) or protease inhibitors (PI).

5.3.5 Outcomes

Outcomes were pregnancy loss among all pregnant women, and gestational age at delivery, birthweight, and size for gestational age among all live singleton births.
Pregnancy loss was defined as any loss before delivery, with the composite variable including non-viable and ectopic pregnancy as determined by research sonographer; miscarriage defined as loss <28 weeks of a clinically recognised pregnancy (14); stillbirths as fetal death occurring before or during labour/delivery (with 1-minute APGAR score of 0).

Gestational age was defined as completed weeks at delivery, categorised as a preterm (<37), normal (≥37-<42) or post-term (≥42) delivery (15). Preterm delivery was also classified as spontaneous or clinician-initiated. Size for gestational age was defined by birthweight percentiles: birthweights

<10th percentile for gestational age as SGA, between 10th and 90th percentile as appropriate for gestational age (AGA), and >90th percentile as large for gestational age (LGA) (16). Secondary outcomes of interest included late/moderately preterm (32–<37), very preterm (28-<32), extremely preterm (<28) delivery. Infant birth weight is commonly used in LMICs, since gestational age at birth is often unknown; we, thus, included this as a supplementary outcome of interest, categorised as low (<2500 grams), normal (2500–4000 grams) or high (HBW, >4000 grams); and very LBW (VLBW, <1500 grams).

5.3.6 Statistical analysis

5.3.6.1 Gestational age bias analysis

Quantitative bias analysis (QBA) is a sensitivity analysis quantifying the magnitude, direction, and effect of any bias (17). Given concerns around LMP and SFH-based gestational age and their potential impact on our association of interest (11), we performed a QBA to estimate the extent to which misclassification of PTD status due to these measures may have over- or under-estimated the association with HIV status, using well-described methods (17). Women with ultrasound-based gestational age, the gold standard measure, were included in an internal validation dataset if they also had a gestational age estimated by either LMP and/or SFH. Preterm delivery variables based on each gestational age assessment method were created, using estimated gestational age at enrolment and date of delivery, to inform this bias analysis.

To address misclassification bias, QBA sensitivity and specificity of the LMP-/SFH-based were calculated in relation to the gold-standard classified ultrasound-based PTD status (Table S1, Part A). These bias parameters are then used to calculate the expected number of PTD cases in each cell of the contingency table (Table S1, Part B); enabling the calculation of measures of association corrected for misclassification bias using the specified equations (Table S1, Part C). We tested three correction strategies for any LMP- and SFH-based outcome misclassification using the overall bias parameters and strata specific bias parameters (assuming non-differential and differential misclassification) (18). To explore the impact of a range of misclassification probabilities on our association of interest we undertook probabilistic bias analysis (PBA), which enables values assigned to the bias parameter to be varied (17). Detailed descriptions of the QBA analyses can be found in Supplemental File 9.3.1.

The association between HIV status and PTD was estimated using ultrasound-based gestational age to determine the “true” PTD status; and also using LMP- and SFH-based gestational age. The magnitude and extent of misclassification bias was assessed by comparing the LMP and SFH risk ratio distribution with the “true” risk ratio (18). To enable accounting for misclassification bias while

using multivariable models, gestational age measurement error was corrected using a multiple overimputation (12). Mismeasured gestational ages were replaced (overimputed) based on observed covariates; mismeasured values were used to create observational level Bayesian priors and variance (due to measurement error) for the imputation model. Five imputation datasets were created, and regression models run for each dataset individually, with the estimated model parameters pooled and summarised (19, 20). Bias corrected gestational age measures were used for the categorisation of PTD and SGA status. Details of multiple overimputation can be found in Supplemental File 9.3.2.

Multiple overimputation and pooling of imputed datasets was performed using the Amelia II and Zelig packages in R (R Foundation for Statistical Computing, Vienna, Austria); all other analyses used STATA version 16.0 (Stata Corporation, College Station, Texas, USA). HIV-negative women and WLHIV were compared; among WLHIV those initiating ART preconception were compared to those initiating ART during pregnancy by trimester. Proportions were compared using chi-squared tests. Associations between HIV status and ART initiation time with adverse birth outcomes were examined using modified Poisson regression with robust error variance, adjusted for maternal age, parity, hypertension and previous PTD as confounders identified *a priori*.

5.3.6.2 Missing data

To account for missing data (>5%), we conducted a sensitivity analysis where we multiply imputed (n=50 imputations) with predictive mean matching for continuous variables (height (8%), BMI (9%), haemoglobin (21%)) and Bayesian polytomous regression for categorical variables (pregnancy outcome (6%)) (21).

5.3.6.3 Sensitivity analysis

To reduce dependency between gestational duration and treatment and ensure all women initiating ART during pregnancy had equal opportunities of experiencing a PTD, only women initiating <35 weeks were included.

5.3.6.4 Ethical approval

Ethical approval was provided by University of Cape Town Faculty of Health Sciences Human Research Ethics Committee and University of Southampton Faculty of Medicine Ethics Committee. Written informed consent was obtained from all women.

5.4 Results

Of 3952 enrolled women, 2507 were HIV-negative (63%) and 1445 living with HIV (37%), of whom 717 (50%) initiated ART before and 728 (50%) during pregnancy (Figure 5.1).

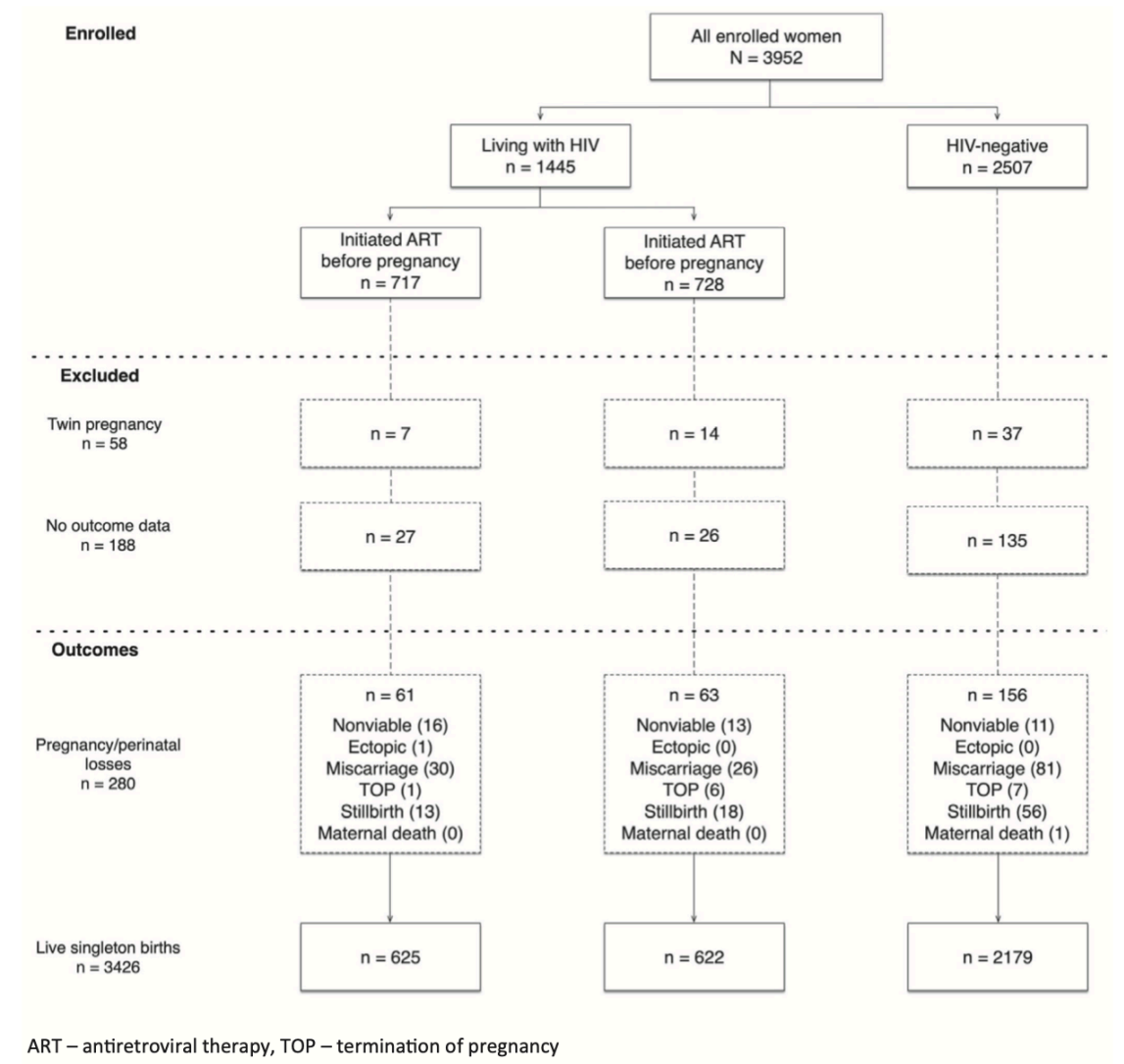

Figure 5.1 Birth outcomes by HIV status and timing of antiretroviral therapy initiation among women in the cohort

ART – antiretroviral therapy, TOP – termination of pregnancy

Women living with HIV were older and less likely to be primigravid than HIV-negative women, as were those who initiated ART before pregnancy compared to those initiating during pregnancy (Table 5.1). ART regimen was nearly always TDF+FTC+EFV, only 2% of women used a PI regimen. Results of imputed baseline variables can be found in Table 9.3.2.

Table 5.1 Baseline characteristics of all enrolled pregnant women at 1st antenatal care visit (n=3952)

	WLHIV *N* = 1445	HIV-negative *N* = 2507	WLHIV *N* = 1445	
			Initiated ART before pregnancy N = 717	Initiated ART during pregnancy N = 728
Age (years)				
≤24	285 (20)	986 (39)	112 (16)	173 (24)
25–29	405 (28)	699 (28)	156 (22)	249 (34)
≥30	755 (52)	822 (33)	449 (63)	306 (42)
Median (IQR)	30 (25, 33)	26 (22, 31)	31 (27, 35)	28 (25, 32)
Height (cm)				
≤155	466 (32)	724 (29)	223 (31)	243 (33)
156–161	487 (34)	882 (35)	251 (35)	236 (32)
≥162	402 (28)	663 (26)	196 (27)	206 (28)
Missing	*90 (6)*	*238 (10)*	*47 (7)*	*43 (11)*
Median (IQR)	158 (154, 162)	158 (154, 162)	158 (154, 162)	158 (154, 162)
Body Mass Index (kg/m^2)				
Normal (≤24.9)	283 (20)	473 (19)	133 (19)	150 (21)
Overweight (25.0–29.9)	393 (27)	627 (25)	205 (29)	188 (26)
Obese (≥30.0)	661 (46)	1162 (46)	324 (45)	337 (46)
Missing	*108 (8)*	*245 (10)*	*55 (8)*	*53 (7)*
Median (IQR)	30 (26, 35)	30 (26, 35)	30 (26, 35)	30 (25, 35)
Gravidity				
1	239 (17)	725 (29)	99 (14)	140 (19)
2	502 (35)	854 (34)	239 (33)	263 (36)
≥3	682 (47)	893 (36)	370 (52)	312 (43)
Missing	*22 (2)*	*35 (1)*	*9 (1)*	*13 (2)*
Parity				
0	327 (23)	866 (35)	146 (21)	181 (25)
1	567 (39)	900 (36)	274 (38)	293 (40)
≥2	528 (37)	706 (28)	288 (40)	240 (33)
Missing	*23 (2)*	*35 (1)*	*9 (1)*	*14 (3)*
Previous preterm				
Yes	125 (9)	162 (6)	65 (9)	60 (8)
Haemoglobin (g/dl)				
Normal (≥11.0)	579 (40)	1122 (45)	340 (47)	239 (33)
Mild anaemia (9.0–10.9)	473 (33)	694 (28)	194 (27)	279 (38)
Moderate anaemia (7.0–8.9)	134 (9)	116 (5)	57 (8)	77 (11)
Severe anaemia (<7.0)	1 (0.1)	4 (0.2)	0	1 (0.1)
Missing	*258 (18)*	*571 (23)*	*127 (18)*	*132 (18)*
Gestational age at 1st antenatal care visit* (weeks)				
Median, (IQR)	18 (12, 24)	19 (14, 25)	16 (12, 24)	19 (13, 25)
Booking blood pressure (mmHG) (%)				
Normal (Systolic <120 or Diastolic <80)	891 (62)	1584 (63)	448 (63)	443 (61)
Prehypertensive (Systolic 120–139 and Diastolic 80–89)	450 (31)	773 (31)	234 (33)	216 (30)
Hypertensive (Systolic >140 or Diastolic >90)	68 (5)	97 (4)	26 (4)	42 (6)
Missing	*36 (2)*	*53 (2)*	*9 (1)*	*27 (4)*
Any hypertension**	145 (10)	235 (9)	71 (10)	74 (10)
Current ART regimen				
TDF–3TC-EFV	1370 (95)	–	638 (89)	667 (92)
Other NNRTI-based regimen	41 (3)	–	33 (5)	8 (1)
PI-based regimen	22 (1)	–	22 (3)	0
Missing	*12 (1)*	–	*24 (3)*	*53 (7)*

ART – antiretroviral therapy, WLHIV – women living with HIV

* best available measure: ultrasound used when available followed by symphysis fundal height then bias-corrected last menstrual period
** hypertension at any point in pregnancy

5.4.1 Gestational age assessment bias analysis

Overall, 88% (n=3475) of women gestational age estimated by LMP, 62% (n=2433) by SFH, and 59% (n=2340) US (Table 9.3.1). Estimated median gestational age at the 1st ANC visit varied by assessment method; however within each assessment method differences were limited to differences of up to one week by HIV status (Suppl. Table 9.3.3). Women with ultrasound-based gestational age did not differ substantially in the variables available from those without ultrasound-based assessment, any unmeasured differences were likely driven by their reasons for attending ANC late (Suppl. Table 9.3.4). Estimated incidence of PTD was 36% based on LMP derived gestational age, on 20% based on SFH gestational age, and 15% based on ultrasound measured gestational age. Estimated PTD incidence was similar for HIV-negative, preconception or initiating ART women using SFH or ultrasound gestational age, with a lack of association with HIV status. Using LMP gestational age, PTD was more common in WLHIV initiating during pregnancy than HIV-negative women or WLHIV initiating preconception (Suppl. Table 9.3.5).

In bias analyses, the observed associations between HIV status and PTD based on SFH and ultrasound were weak, for both these measure. After adjustment for misclassification by QBA, using overall and non-differential strata-specific bias parameters, there was a slight change in SFH estimates (Suppl. Table 9.3.6), suggesting the SFH gestational age measure did not produce a biased association. In contrast, the observed association between HIV status and PTD using LMP gestational age showed increased PTD risk for WLHIV. Following covariate adjustment, with all three misclassification scenarios final estimates were biased in either direction in relation to the null, with non-differential strata-specific bias parameters producing the most biased estimates (Suppl. Table 9.3.6). Given the magnitude of bias observed in the estimates using LMP gestational age, multiple overimputation was used for LMP gestational age, with the overimputed gestational age used in birth outcome analyses.

5.4.2 Adverse birth outcomes by HIV status and timing of ART initiation

3706 women were included in the pregnancy loss analysis (Figure 5.1). Eight percent (n=280) of pregnancies ended in a loss: nonviable pregnancy (14%), miscarriages (49%) and stillbirths (31%) (Figure 5.1; Table 5.2). In adjusted analyses, there was a suggested modest increase in risk by HIV status (adjusted risk ratio (ARR) 1.26, 95% CI 0.98, 1.61); risk was similar by timing of ART initiation (ARR 1.09, 95% CI 0.77, 1.54).

Table 5.2 Birth Outcomes by HIV status and timing of antiretroviral therapy initiation among enrolled women with singleton pregnancies (n=3894)

	WLHIV *N* = 1424	HIV-negative *N* = 2470	WLHIV N = 1424	
			Initiated ART before pregnancy N = 710	Initiated ART during pregnancy N = 714
Pregnancy outcome				
Pregnancy loss	124 (8)	156 (6)	61 (9)	63 (9)
Miscarriage	56 (4)	81 (3)	30 (4)	26 (4)
Stillbirth	31 (2)	56 (2)	13 (2)	18 (2)
Nonviable	29 (2)	11 (0.4)	16 (2)	13 (2)
Other[ξ]	8 (0.6)	8 (0.3)	2 (0.3)	6 (1)
Live Births	1247 (88)	2179 (88)	622 (88)	625 (88)
Missing	*53 (4)*	*135 (6)*	*27 (4)*	*26 (4)*
Gestational length (weeks)				
Post-term (≥42)	43 (3)	134 (6)	23 (4)	20 (3)
Term (37–41)	979 (79)	1677 (77)	491(79)	488 (78)
Preterm (<37)*	225 (18)	368 (17)	108 (17)	117 (19)
Late/moderate preterm (32-<37)	182 (15)	304 (14)	84 (14)	98 (16)
Very preterm (28–31)	32 (2)	48 (2)	19 (3)	13 (2)
Extremely preterm (<28)	9 (0.8)	12 (0.6)	3 (0.5)	6 (1)
Clinical presentation at labour				
Spontaneous	118 (9)	205 (9)	54 (9)	64 (10)
Clinician-initiated	96 (8)	147 (7)	48 (7)	48 (8)
Unknown	11 (1)	16 (1)	6 (1)	5 (1)
Missing	*16 (1)*	*39 (2)*	*6 (1)*	*10 (2)*
Size for gestational age (centile)				
LGA (>90th)	175 (14)	350 (16)	82 (13)	93 (15)
AGA (10th – 90th)	883 (71)	1508 (69)	442 (71)	441 (71)
SGA (<10th)	137 (11)	169 (8)	72 (12)	65 (10)
Missing	*52 (4)*	*152 (7)*	*26 (4)*	*26 (4)*

AGA – appropriate for gestational age, ART – antiretroviral therapy, LGA – large for gestational age, SGA – small for gestational age, WLHIV – women living with HIV
ξ other losses include abortion (n=14), ectopic pregnancy (n=1) and maternal death (n=1)
* preterm deliveries not classified because there was no defined gestational age (n=6)

Following exclusions 3426 women were included in the PTD and SGA analyses (Figure 5.1). Overall, 78% (n=2656) of deliveries were term, 17% preterm and 5% post-term, with little difference by HIV status and ART initiation timing (Table 5.2). Adjusted PTD risk was similar by HIV status (ARR 1.02, 95% CI 0.89, 1.20) (Table 5.3, comparison A), and by ART status among WLHIV (ARR 1.12, 95% CI 0.88, 1.43) (Table 5.3, comparison B). Compared to HIV-negative women, PTD risk in WLHIV with preconception ART and those initiating during pregnancy was similar (Table 5.3, comparison C). Among WLHIV there appeared to be a trend of increasing PTD risk by trimester, with highest risk in the third trimester (Table 5.3, comparison D; Suppl. Table 9.3.7). Women with longer gestations have increased opportunity to initiate ART treatment, but in analysis restricted to women initiating <35 weeks gestational age, results did not change. Of the 593 PTD cases, 82% (n=486) were late/moderate, 13% very preterm (n=80) and 4% extremely preterm delivery (n=21), with little difference by HIV and ART initiation timing status (Table 5.2).

Overall 9% of infants were classified as SGA, higher in WLHIV (11%) than HIV-negative women (8%) (Table 5.2). In adjusted analyses, infants of WLHIV were at increased risk of being SGA (ARR 1.37,

95% CI 1.10, 1.72) (Table 5.3, comparison A), both for preconception and during pregnancy ART initiators (Table 5.3, comparison C). Among initiators of ART during pregnancy, infants of first trimester initiators had a similar risk of being SGA to HIV-negative women; while second trimester initiators were at increased risk of having SGA infants compared to HIV-negative women (Suppl. Table 9.3.4; Table 5.3, comparison D). The lowest proportion of SGA infants was seen among third trimester initiators (Suppl. Table 9.3.7).

Overall, 13% of infants were LBW; slightly higher among WLHIV (14%) than HIV-negative (12%) women. No differences were observed by HIV status or timing of ART initiation (Suppl. Table 9.3.9 and 9.3.10, Supplemental File 9.3.3).

Table 5.3 Associations between HIV and ART exposure groups and adverse birth outcomes

Outcome measure	Comparison A†	RR (95% CI)	ARR (95% CI)	Comparison B††	RR (95% CI)	ARR (95% CI)	Comparison C†††	RR (95% CI)	ARR (95% CI)	Comparison D††††	RR (95% CI)	ARR (95% CI)
Preterm delivery (<37 weeks)*	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.06 (0.9	1.02 [illegible]0)	During pregnancy	1.09 (0.84, 1.41)	1.12 (0.88, 1.43)	Before pregnancy	1.02 (0.84, 1.24)	0.97 (0.79, 1.18)	Before pregnancy	1.02 (0.82, 1.26)	0.97 (0.79, 1.18)
										1st trimester	1.05 (0.72, 1.54)	1.01 (0.71, 1.44)
							During pregnancy	1.11 (0.92, 1.34)	1.07 (0.89, 1.31)	2nd trimester	1.07 (0.82, 1.40)	1.05 (0.82, 1.34)
										3rd trimester	1.28 (0.86, 1.91)	1.26 (0.89, 1.79)
Small for gestational age (<10th centile)**	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.42 (1.13, 1.77)	1.37 (1.10, 1.72)	During pregnancy	0.90 (0.64, 1.26)	0.96 (0.69, 1.34)	Before pregnancy	1.49 (1.13, 1.97)	1.39 (1.05, 1.84)	Before pregnancy	1.49 (1.13, 1.97)	1.39 (1.05, 1.84)
										1st trimester	1.31 (0.79, 2.19)	1.25 (0.77, 2.03)
							During pregnancy	1.34 (1.01, 1.78)	1.36 (1.03, 1.78)	2nd trimester	1.54 (1.09, 2.16)	1.53 (1.11, 2.12)
										3rd trimester	0.87 (0.43, 1.78)	0.98 (0.49, 1.93)

WLHIV – women living with HIV

†Comparison A (WLHIV vs HIV-negative), ††Comparison B (WLHIV ART initiated during pregnancy vs WLHIV ART initiated before conception), †††Comparison C (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy vs HIV-negative), ††††Comparison D (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy (1st trimester) vs WLHIV ART initiated during pregnancy (2nd trimester) vs WLHIV ART initiated during pregnancy (3rd trimester) vs HIV-negative)

* adjusted for age, preeclampsia, parity, hypertension

** adjusted for age, maternal height, hypertension, BMI, parity, preeclampsia

5.5 Discussion

5.5.1 Principal findings

We evaluated birth outcomes in a routine care cohort of women seeking ANC at a primary care facility, using multiple overimputation to correct for measurement error in gestational age. Compared to HIV-negative women, there was no appreciable difference in risk of PTD in WLHIV on ART from before or during pregnancy; however, WLHIV were at increased risk of having an SGA infant (regardless of ART initiation timing). While these results are reassuring and indicate relative safety of the TDF+FTC+EFV regimen, in this treat-all era, with increasing treatment duration and ART regimen options, optimisation of HIV treatment choices for pregnant and reproductive age women should remain a priority.

5.5.2 Strengths of the study

Strengths of this study include recruitment of a large community-based cohort. Accurate gestational age is essential in birth outcomes studies, as use of imprecise methods can affect quantification of associations (11). Use of a research sonographer for gestational age assessment in women ≤24 weeks when ultrasound is highly reproducible and accurate for dating is another strength. In women with less reliable gestational age measures, our analysis was strengthened by the use of QBA to give an estimate of the magnitude, direction and uncertainty arising from gestational age assessment by SFH and, especially, LMP, and subsequently using multiple overimputation to address measurement error. Additionally, our study provides quantification of adverse birth outcome risk, and insights into potential mechanisms underlying outcomes.

5.5.3 Limitations of the data

Although routinely collected data was used to confirm birth outcomes, several other variables were self-reported, which could have introduced measurement error. Because recruitment occurred at the 1st ANC visit, earlier pregnancy losses were not documented. There may be potential residual confounding of our results due to unmeasured factors: ultrasound was available for women assessed as being ≤24 weeks by LMP/SFH, and due to limited socio-demographic data, differences between those attending care before or after that cut-off could not be explored in detail. Further, we only had basic maternal characteristics for use in adjusted analyses of the impact of timing of ART on birth outcomes; we lacked detailed data on other potential confounding factors such as maternal nutritional status, substance use, ART adherence and mental health. We were unable to address immortal time bias using a “time-varying” approach to allow for actual timing of ART use because we did not have an untreated WLHIV cohort. Lastly, few women received a PI-based ART

regimen, more commonly initiated in advanced HIV, and our sample size was thus insufficient to detect increased PTD risk associated with PI.

5.5.4 Interpretation

Our finding of a similar frequency of PTD in WLHIV and HIV-negative women is consistent with recent results in similar settings (22), including one study with an ultrasound examination at ≤24 weeks (23). Conversely, other studies have reported an increased risk of PTD in WLHIV (6, 10, 24); including a previous analysis in the same clinic using ultrasound-based gestational age, where WLHIV were twice as likely to deliver preterm (25). Differences between studies have been suggested to be at least partially due to exposure to specific ART regimens (26). Increasingly safer and efficacious regimens have been introduced, initiated at or soon after HIV diagnosis while still asymptomatic, and WLHIV initiating preconception are now likely to be using the same regimen as those initiating during pregnancy. This can be seen when comparing regimens in our previous and current analysis; previously approximately 34% of WLHIV initiating ART preconception were on TDF+FTC+EFV (25), compared with 90% currently. Improved immunologic status of WLHIV as a consequence of higher and earlier uptake of ART could explain the lack of differences by HIV status seen here. Evidence of this improved immunologic status could possibly be demonstrated by comparing PTD incidence in our previous (23%) analysis, when early treatment was less widespread (25), and our current analysis (18%) during the treat-all era. Support for this suggestion also comes from other studies (10, 27), and a study showing an improvement in birth outcomes with increased coverage of TDF+FTC+EFV (6).

WLHIV had increased risk of SGA infants compared to HIV-negative women. Although SGA may be due to advanced HIV infection (28), studies from the ART era (5, 9) show treated WLHIV to be at increased risk of SGA infants compared to HIV-negative women, suggesting ART is also involved. Our finding that women initiating before or in early pregnancy were at increased risk of having SGA infants, could point to placenta-mediated mechanisms. Fetal growth and development depend on robust placental functioning (29), with maternal infections able to disrupt placenta development (30). In a subset of HIV-infected women in PIMS with placentas collected (13), we found substantially increased maternal vascular malperfusion in placentas from women initiating ART preconception, suggestive of placental dysfunction, in turn associated with adverse birth outcomes (31). Maternal HIV infection with ART exposure has also been associated with reduced placental growth factor levels and a subsequent increased SGA risk (32). We hypothesise that the increased SGA risk we observed was due to placental functioning compromised by preconception or early ART

exposure through disruption of vasculogenesis and angiogenesis, with dysregulation of pro-angiogenic mediators.

Pregnancy is associated with increased immune activation (33), while ART aims to reduce HIV-related immune activation, which in pregnancy may potentially alter regulatory pathways required for a healthy pregnancy (34). Our initial hypothesis that ART initiation before pregnancy would be associated with PTD and during pregnancy with SGA risk was not confirmed, and we did not observe differences in outcomes by ART initiation timing. This differed from a meta-analysis (35) which showed increased PTD and LBW risk among women on preconception ART, but no differences in SGA by timing of initiation. A key confounder in the relationship between ART initiation timing and adverse outcomes is maternal disease stage; previously women initiating ART preconception-initiated treatment because of advanced HIV disease status. In the era of our cohort, ART had become common from before or early pregnancy; we showed in a case-control subset of women in our PIMS cohort that early and prolonged ART use may lead to correction of the immediate immunological response following ART initiation (36).

5.6 Conclusion

Using quantitative bias analyses and multiple overimputation to detect and correct for measurement error in gestational age, we show increased ART-associated SGA risk but not increased PTD risk. Mechanistic understanding of the contribution of ART to adverse birth outcomes needs to be improved to optimise treatment for pregnant WLHIV and ensure optimal maternal and infant outcomes. Monitoring potential risks associated with HIV status and ART initiation timing in a population with very high HIV levels remains required, especially with the introduction of new antiretroviral agents.

5.7 References

1. WHO. Global Action Plan: For health lives and well-being for all. Geneva, Switzerland: World Health Organization; 2018.
2. Katz J, Lee AC, Kozuki N, et al. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *Lancet*. 2013;382(9890):417-425.
3. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Geneva, Switzerland: World Health Organization; 2016.
4. UN-IGME. Levels and Trends in Child Mortality Report 2018. New York, USA: United Nations Inter-agency group for child mortality estimation; 2018.
5. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.
6. Moodley T, Moodley D, Sebitloane M, Maharaj N, Sartorius B. Improved pregnancy outcomes with increasing antiretroviral coverage in South Africa. *BMC Pregnancy Childbirth*. 2016;16(1):35-45.
7. Li N, Sando MM, Spiegelman D, et al. Antiretroviral therapy in relation to birth outcomes among HIV-infected women: a cohort study. *J Infect Dis.* 2016;213(7):1057-1064.
8. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.
9. Li H, Liu J, Tan D, et al. Maternal HIV infection and risk of adverse pregnancy outcomes in Hunan province, China: A prospective cohort study. *Medicine (Baltimore)*. 2020;99(8):e19213
10. Zash R, Souda S, Chen JY, et al. Reassuring Birth Outcomes with Tenofovir/Emtricitabine/ Efavirenz used for Prevention of Mother to Child Transmission of HIV in Botswana. *J Acquir Immune Defic Syndr.* 2016;71(4):428-436.
11. Malaba TR, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol.* 2018;28(12):893-900.
12. Bengtson AM, Westreich D, Musonda P, et al. Multiple overimputation to address missing data and measurement error: application to HIV treatment during pregnancy and pregnancy outcomes. *Epidemiology*. 2016;27(5):642-650.
13. Malaba TR, Gray CM, Myer L, Newell ML. Cohort Profile: Prematurity Immunology in HIV-infected Mothers and their infants Study (PIMS). *BMJ Open*. 2021;11(9):e047133

14. WHO. Neonatal and perinatal mortality: country, regional and global estimates. Geneva, Switzerland: World Health Organization; 2006.
15. WHO. Preterm Birth. Geneva, Switzerland: World Health Organization; 2017.
16. Villar J, Ismail LC, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet.* 2014;384(9946):857-868.
17. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Springer Science & Business Media: Berlin, Germany; 2011.
18. Walraven CV. A comparison of methods to correct for misclassification bias from administrative database diagnostic codes. *Int J Epidemiol.* 2018 Apr 1;47(2):605-616.
19. Choirat C, Honaker J, Imai K, King G, Lau O. Zelig: Everyone's statistical software. 2017, Version 5.0–15; 2017.
20. Imai K, King G, Lau O. Toward a common framework for statistical analysis and development. *J Comput Graph Stat*. 2008;17(4):892-913.
21. Rubin D. Multiple imputation for nonresponse in surveys. John Wiley & Sons: New York, USA 1987.
22. González R, Rupérez M, Sevene E, et al. Effects of HIV infection on maternal and neonatal health in southern Mozambique: A prospective cohort study after a decade of antiretroviral drugs roll out. *PLoS One*. 2017;12(6):e0178134.
23. Goldman G, Budhram S. A retrospective cohort study comparing pregnancy outcomes and neonatal characteristics between HIV-infected and HIV-non-infected mothers. *S Afr Med J*. 2020;110(6):502-504.
24. Zash R, Jacobson DL, Diseko M, et al. Comparative safety of antiretroviral treatment regimens in pregnancy. *JAMA Pediatr*. 2017;171(10):e172222.
25. Malaba TR, Phillips TK, Le Roux S, et al. Antiretroviral therapy use during pregnancy and adverse birth outcomes in South African women. *Int J Epidemiol*. 2017;46(5):1678-1689.
26. Mofenson LM. Antiretroviral therapy and adverse pregnancy outcome: the elephant in the room? *J Infect Dis*. 2016;213(7):1051-1054.
27. Mehta UC, Van Schalkwyk C, Naidoo P, et al. Birth outcomes following antiretroviral exposure during pregnancy: Initial results from a pregnancy exposure registry in South Africa. *South Afr J HIV Med*. 2019;20(1):971-983.
28. Ndirangu J, Newell ML, Bland RM, Thorne C. Maternal HIV infection associated with small-for-gestational age infants but not preterm births: evidence from rural South Africa. *Hum Reprod*. 2012;27(6):1846-1856.

29. Geva E, Ginzinger DG, Zaloudek CJ, et al. Human placental vascular development: vasculogenic and angiogenic (branching and nonbranching) transformation is regulated by vascular endothelial growth factor-A, angiopoietin-1, and angiopoietin-2. *J Clin Endocrinol Metab.* 2002;87(9):4213-4224.
30. Weckman AM, Ngai M, Wright J, McDonald CR, Kain KC. The impact of infection in pregnancy on placental vascular development and adverse birth outcomes. *Front Microbiol.* 2019;10:1924-1935.
31. Ikumi NM, Malaba TR, Pillay K, et al. Differential impact of antiretroviral therapy initiated before or during pregnancy on placenta pathology in HIV-positive women. *AIDS.* 2021;35(5):717-726.
32. Conroy AL, McDonald CR, Gamble JL, et al. Altered angiogenesis as a common mechanism underlying preterm birth, small for gestational age, and stillbirth in women living with HIV. *Am J Obstet Gynecol.* 2017;217(6):684. e1- e17.
33. Shah NM, Herasimtschuk AA, Boasso A, et al. Changes in T cell and dendritic cell phenotype from mid to late pregnancy are indicative of a shift from immune tolerance to immune activation. *Front Immunol*. 2017;8(1138):1-14.
34. Hunt PW, Sinclair E, Rodriguez B, et al. Gut epithelial barrier dysfunction and innate immune activation predict mortality in treated HIV infection. *J Infect Dis.* 2014;210(8):1228-1238.
35. Uthman OA, Nachega JB, Anderson J, et al. Timing of initiation of antiretroviral therapy and adverse pregnancy outcomes: a systematic review and meta-analysis. *Lancet HIV.* 2017;4(1):e21-e30.
36. Mdletshe N, Thobakgale C, Malaba TR, et al. Low immune activation in early pregnancy is associated with preterm but not small-for-gestational age delivery in HIV infected women initiating antiretroviral therapy in pregnancy: a PIMS case-control study in Cape Town, South Africa. *Clin Infect Dis*. 2021;73(12):2205-2216.

Chapter 6: Blood pressure trajectories during pregnancy and associations with adverse birth outcomes among HIV-infected and HIV-uninfected women in South Africa: a group-based trajectory modelling approach.

Malaba TR, Cois A, Madlala HP, Matjila M, Myer L, Newell ML. Blood pressure trajectories during pregnancy and associations with adverse birth outcomes among HIV-infected and HIV-uninfected women in South Africa: a group-based trajectory modelling approach. *BMC Pregnancy Childbirth*. 2020;20(1):742-754.

Relevance of this paper to the thesis

This chapter presents data on blood pressure assessed longitudinally over pregnancy, as high blood pressure late in pregnancy has been associated with adverse birth outcomes and is also associated with HIV and antiretroviral therapy. This chapter presents one of the first exploration of discrete latent class assignment of blood pressure trajectory groups during pregnancy and their relationship with adverse birth outcomes, in a population-based study. There is also investigation of whether blood pressure modifies the association between maternal HIV and antiretroviral therapy and adverse birth outcomes.

Contribution of the student and co-authors

TRM conceptualised and conducted all analyses, led data interpretation and drafted the manuscript. TRM also directed data collection, with assistance from HM. MLN and AC provided input into the data analysis and assisted with data interpretation. MLN and LM conceptualised the study from which these data arise, were responsible for funding, implementation and overall leadership. All authors reviewed and approved the final manuscript.

6.1 Abstract

6.1.1 Background

High blood pressure (BP) late in pregnancy is associated with preterm delivery (PTD); BP has also been associated with HIV and antiretroviral therapy (ART), but whether the relationship between BP assessed longitudinally over pregnancy and PTD and low birthweight (LBW) is modified by HIV/ART is unclear. We hypothesise the presence of distinctive BP trajectories and their association with adverse birth outcomes may be mediated by HIV/ART status.

6.1.2 Methods

We recruited pregnant women at a large primary care facility in Cape Town. BP was measured throughout pregnancy using automated monitors. Group-based trajectory modelling in women with ≥3 BP measurements identified distinct joint systolic and diastolic BP trajectory groups. Multinomial regression assessed BP trajectory group associations with HIV/ART status, and Poisson regression with robust error variance was used to assess risk of PTD and LBW.

6.1.3 Results

Of the 1583 women in this analysis, 37% were living with HIV. Seven joint trajectory group combinations were identified, which were categorised as normal (50%), low normal (25%), high normal (20%), and abnormal (5%). A higher proportion of women in the low normal group were living with HIV than without HIV (28% vs. 23%), however differences were not statistically significant (RR 1.27, 95% CI 0.98-1.63, reference category: normal). In multivariable analyses, low normal trajectory (ARR0.59, 0.41-0.85) was associated with decreased risk of PTD, while high normal (ARR1.48, 1.12-1.95) and abnormal trajectories (ARR3.18, 2.32-4.37) were associated with increased risk of PTD, and abnormal with increased risk of LBW (RR2.81, 1.90-4.15).

6.1.4 Conclusions

While HIV/ART did not appear to mediate the BP trajectories and adverse birth outcomes association, they did provide more detailed insights into the relationship between BP, PTD and LBW for women living with and without HIV.

6.2 Introduction

Highly elevated blood pressure (BP) levels during pregnancy affect 10-15% of all pregnancies (1, 2). Hypertensive disorders in pregnancy (HDP), including gestational and chronic hypertension, pre-eclampsia and eclampsia, are associated with increased risk of maternal and neonatal morbidity and mortality (1, 3, 4); specifically increased risk of adverse pregnancy outcomes including fetoplacental insufficiency, preterm delivery (PTD) and low birth weight (LBW) infants (5, 6). Differential HDP risk has been theorised to exist in immunocompromised states, such as HIV-infection in pregnancy especially with antiretroviral therapy (ART) use. This was demonstrated in a recent systematic review showing increased HDP risk in women living with HIV on ART, particularly those using protease inhibitors (PI) based regimens, compared to untreated women living with HIV (7). In addition to concerns around increased HDP risk, both untreated advanced HIV infection (8, 9) and ART use (10-12) have also been associated with increased risk of adverse pregnancy outcomes.

Accurate blood pressure measurement is vital to identify HDP, guiding diagnosis, antihypertensive treatment and timing of delivery. Clinically, BP is routinely measured throughout the antenatal period (13), with the expectation that BP will decrease steadily until mid-pregnancy and then increase slightly again until delivery (30). However, in population-based research studies BP is typically analysed using discrete, cross-sectional assessments at arbitrary time points (such as at the first antenatal care (ANC) or enrolment visit) (14), with high BP defined by cut-off points used in non-pregnant adults. Characterising BP trajectories during pregnancy in population-based studies may better inform understanding of the mechanisms driving BP changes in pregnancy and development of HDP, and their timing, than cross-sectional assessments. BP trajectories may also provide more accurate insights into the relationship between BP and maternal, perinatal and neonatal outcomes than cross-sectional assessments and could thus inform development of interventions for improvement in HDP management.

Elucidating the HIV/ART and HDP relationship is critical to understanding the etiology of HDP; however studies investigating this association have generally used traditional hypertension cut-offs from BP readings at first ANC visit. Contrary to the use of serial BP measurements in clinical care, few studies have assessed BP trajectories across pregnancy to identify women at risk of developing HDP and/or adverse pregnancy outcomes. We hypothesised that within a cohort of pregnant women there are groups of individuals that follow distinctive BP trajectories that may not be identifiable using traditional cross-sectional hypertension cut-offs; and that these may differ according to HIV/ART status. The aim of this analysis was to describe the natural history of blood

pressure throughout pregnancy in women living with and without HIV and assess whether this mediates the association between ART use and adverse pregnancy outcomes.

6.3 Methods

6.3.1 Study Population

Between April 2014 and October 2016, consecutive pregnant women (aged ≥18 years), regardless of their HIV status, making their first ANC visit at a large public sector primary care facility in Cape Town, South Africa were enrolled in a prospective cohort (15).

As part of routine ANC services all women had systolic (SBP) and diastolic (DBP) blood pressure measured using automated monitors (Edan M3A Vital Signs Monitor), at the first and subsequent ANC visits. For BP, there was a minimum of 5 minutes rest before a single measurement took place and women were seated upright with the arm supported and the use of appropriately sized cuffs for arm circumference (25-35 cm and 35.5-46 cm cuffs used). BP and GA were measured by regularly trained ANC midwives and recorded in the clinical records. Gestational age (GA) was routinely established at the first ANC visit based on last menstrual period (LMP) and symphysis-fundal height (SFH), with any woman clinically assessed (LMP and/or SFH) to be ≤24 weeks gestation referred for a same day ultrasound by the PIMS research sonographer using standardised assessment protocols and blinded to the midwife GA assessment.

Women living with HIV all on ART either initiated ART pre-conception, continuing their current regimen throughout pregnancy, or initiated ART during pregnancy; with the predominant regimen being a fixed-dose combination of TDF+FTC+EFV.

6.3.2 Data Collection

Following the first ANC visit, data from enrolled women were abstracted from clinical records including HIV status, pregnancy history (previous and index pregnancies), medications prescribed, and any maternal diagnoses during pregnancy. Following delivery, all BP recordings during pregnancy (including measurement date) and obstetric outcomes (including date of delivery and birthweight), were abstracted from clinical records. To avoid inclusion of spuriously high values, all measurements occurring during labour in inpatient settings were excluded because BP medications may be temporarily withheld or because labour-related stress or other emergent conditions may cause temporary elevations in BP. Because of repeated BP measurements during acute and unstable periods throughout pregnancy, 13 GA time point windows were set up with intervals corresponding with ANC visit intervals (4 week intervals until 32 weeks GA after which 2 week intervals until 41

completed weeks). Within each of the 13 GA time point windows, the BP measurement was allocated based on the first available measurement, discarding any remainder in that window.

6.3.3 Outcomes

Prehypertension was defined as BP readings with SBP from 120 to 139mmHg and/or DBP from 80 to 89 mmHg; hypertension was defined as SBP ≥140 mmHg and/or diastolic pressure ≥90 mmHg on more than two occasions (16) and classified according to gestational age at onset. Hypertension classification was according to ISSHP guidelines (17) but based only on BP measurements (i.e. not taking into account proteinuria). BP was classified as chronic if assessed <20 weeks GA (assuming this existed pre-conception); and as gestational when initially normotensive at <20 weeks and then hypertensive >20 weeks or assessed for the first time >20 weeks hypertension. Non-severe hypertension was defined as SBP from 140-159 mmHg and/or DBP from 90-109 mmHg) and severe hypertension defined as SBP ≥160 and/or DBP ≥110 mmHg). Mean arterial pressure (MAP) was calculated using the formula: MAP = (SBP + 2 x DBP)/3 (18).

Secondary outcomes included PTD and LBW. PTD was defined as delivery at <37 weeks' gestation, categorised as late preterm (34–37 weeks), moderately preterm (32-34 weeks) or very preterm (<32 weeks), as assessed by the most reliable available GA assessment method at baseline (ultrasound, SFH or LMP). LBW was defined as birthweight <2500 grams and very low birth weight (VLBW) as <1500 grams. Pregnancies not resulting in the delivery of a live born infant were defined as a pregnancy loss.

6.3.4 Statistical Analyses

Sample characteristics are described as median and interquartile range for numerical variables and frequency for categorical measures. Group-based trajectory modelling (GBTM) was used to analyse changes in participants' SBP and DBP during pregnancy (19). GBTM is an application of finite mixture modelling, allowing the identification of population subgroups (classes) characterised by statistically distinct trajectories for one or more outcomes of interest. In each class, the outcomes trajectories are summarised by polynomial functions of time (21). Model parameters are calculated by maximising a likelihood function and include estimates of individual probabilities of class membership and of the polynomial coefficients which define the shape of the trajectories for each class (20). In addition to estimates of class membership probability, the joint model also estimates conditional probability linking membership across the trajectory classes of both outcomes of interest (20). We applied GBTM to identify classes of pregnant women characterised by different joint trajectories for SBP and DBP during pregnancy. We chose the optimal number of classes and the

degree of the polynomials used to represent the trajectories by fitting multiple models with different combinations of these parameters and comparing their fit to the data through the Bayesian information criterion (BIC), which balances model fit and complexity, with smaller BIC values indicating a better fit (21). Linear, quadratic and cubic polynomial were tested, and number of classes varying between 1 and 7.

After fitting the chosen model, women were assigned to SBP and DBP classes based on the highest class membership probability. The probability matrix of joint class membership and considerations of clinical significance of the observed differences were then used to assign women to 'joint groups' (groups hereafter) characterised by different combinations of SBP and DBP trajectories, with descriptive labels assigned based on the visual patterns of BP changes throughout pregnancy. Multinomial regression was then used to identify predictors of group membership, and Poisson regression with robust error variance (22), to investigate the association between the group membership and adverse birth outcomes (PTD and LBW).

Statistical analyses were performed using STATA version 14.0 (Stata Corporation, College Station, TX, USA) and its TRAJ plugin to perform GBTM (20). Two-tailed chi-squared and rank-sum tests were used, as appropriate, to compare participants characteristics across baseline categories and trajectory groups. Tests were conducted at the level of significance α=5%.

6.4 Results

A total of 1538 women with at least three antenatal BP readings (median number of measurements 5, interquartile range, IQR 4-8) and a live singleton birth were included in this analysis: 967 living without HIV (63%) and 571 living with HIV (37%); 54% of the latter group initiated ART pre-conception (n=306) and 46% during pregnancy (n=265) (Table 6.1). Women living with HIV were older, less likely to be primigravid and more likely to be anaemic than women living without HIV (Table 6.1). Women living with HIV who initiated ART before pregnancy were older, less likely to be primigravid and anaemic than those who initiated during pregnancy. Overall, 25% of women were overweight and 50% obese, similar by HIV status and timing of ART initiation (Table 6.1).
At the first ANC visit, 31% of women were classified as prehypertensive and 5% hypertensive, with no differences by HIV status. Differences were however observed by ART status with a higher proportion of women initiating ART in pregnancy classified as hypertensive compared to those initiating preconception (7% vs 4%). Using the cross-sectional hypertension assessments at multiple time points antenatally, 22% of women were classified as being hypertensive at least once with no

differences according to HIV/ART status. For all those who were hypertension at any point, 13% were hypertensive <20 weeks GA and thus classified as having chronic hypertension; 87% were either initially assessed to be normotensive at <20 GA and were subsequently hypertensive or were assessed >20 weeks GA as hypertensive and classified as having gestational hypertension (Table 6.1).

Table 6.1 Baseline description of women with at least 3 pre-labour blood pressure readings

	Total ***N* = 1538**	**HIV-uninfected** ***N* = 967**	**HIV-infected** ***N* = 571**	***P* value***	**HIV-infected** ***N* = 571**		***P* value****
					Initiated before pregnancy ***N* = 306**	**Initiated during pregnancy** ***N* = 265**	
Age, years (%)				< 0.0001			< 0.0001
< 24	429 (28)	329 (34)	100 (18)		47 (15)	53 (20)	
25–29	454 (30)	290 (30)	164 (29)		71 (23)	93 (35)	
> 30	655 (43)	348 (36)	307 (54)		188 (61)	119 (45)	
Median (IQR)	28 (24–32)	27 (23–31)	30 (26–34)		31 (27–35)	29 (25–32)	
Height, cm (%)				0.177			0.430
≤ 155	470 (30)	301 (31)	263 (30)		87 (28)	82 (31)	
156–161	533 (35)	341 (35)	291 (33)		102 (33)	90 (34)	
≥ 162	412 (27)	242 (25)	250 (30)		93 (30)	77 (29)	
Missing	*123 (8)*	*141 (9)*	*53 (7)*		*24 (8)*	*16 (6)*	
Median (IQR)	158 (154–162)	158 (154–162)	158 (154–163)		158 (154–163)	158 (155–163)	
Body Mass Index, kg/m^2 (%)				0.074			0.121
Underweight (< 18.5)	9 (0.6)	4 (0.4)	5 (0.9)		3 (1)	2 (0.8)	
Normal (18.5–24.9)	246 (16)	144 (15)	102 (18)		49 (16)	53 (20)	
Overweight (25.0–29.9)	386 (25)	241 (25)	145 (25)		85 (28)	60 (23)	
Moderately Obese (30.0–34.9)	364 (24)	221 (23)	143 (25)		80 (26)	63 (24)	
Severely Obese (> 35.0)	397 (26)	270 (28)	127 (22)		59 (19)	68 (26)	
Missing	*136 (9)*	*87 (9)*	*49 (9)*		*30 (10)*	*19 (7)*	
Median (IQR)	31 (26–36)	31 (27–37)	30 (26–35)		30 (26–34)	31 (26–36)	
Gravidity (%)				0.001			0.001
1	291 (19)	210 (22)	81 (14)		35 (11)	46 (17)	
2	575 (37)	363 (37)	212 (37)		106 (35)	106 (40)	
≥ 3	666 (43)	393 (41)	273 (48)		162 (53)	111 (42)	
Missing	*6 (0.4)*	*1 (0.1)*	*5 (0.9)*		*3 (1)*	*2 (0.8)*	
Median (IQR)	2 (2–3)	2 (2–3)	2 (2–3)		3 (2–3)	2 (2–3)	
Parity (%)				0.004			< 0.0001
0	383 (25)	268 (28)	115 (20)		56 (18)	59 (22)	
1	644 (42)	396 (41)	248 (43)		121 (40)	127 (48)	
≥ 2	505 (33)	302 (31)	203 (36)		126 (41)	77 (29)	
Missing	*6 (0.4)*	*1 (0.1)*	*5 (0.9)*		*3 (1)*	*2 (0.8)*	
Median (IQR)	1 (1–2)	1 (0–2)	1 (1–2)		1 (1–2)	1 (1–2)	
Previous Preterm (%)				0.230			0.368
Yes	127 (8)	74 (8)	53 (9)		26 (9)	27 (10)	

Table 6.1 Baseline description of women with at least 3 pre-labour blood pressure readings cont.

	Total *N* = 1538	HIV-uninfected *N* = 967	HIV-infected *N* = 571	*P* value*	HIV-infected *N* = 571		*P* value**
					Initiated before pregnancy *N* = 306	Initiated during pregnancy *N* = 265	
Haemoglobin g/dl (%)				< 0.0001			< 0.0001
Normal (≥11.0)	567 (37)	369 (38)	199 (35)		119 (39)	80 (30)	
Mild Anaemia (9–10.9)	401 (26)	238 (25)	163 (29)		81 (26)	82 (31)	
Moderate Anaemia (7–8.9)	94 (6)	42 (4)	52 (9)		19 (6)	33 (12)	
Severe Anaemia (< 7)	2 (0.1)	2 (0.2)	0		0	0	
Missing	*474 (32)*	*317 (33)*	*157 (28)*		*87 (28)*	*70 (26)*	
Gestational Age Assessment[a]							
Median (weeks) (IQR)	18 (13–24)	18 (13–24)	18 (12–24)	0.252	17 (12–23)	19 (13–24)	0.334
Booking Blood Pressure (mmHG) (%)				0.967			0.575
Normal (Sys < 120 or Dia < 80)	989 (64)	622 (64)	367 (64)		196 (64)	171 (65)	
Prehypertensive (Sys 120–139 and Dia 80–89)	471 (31)	297 (31)	174 (31)		98 (32)	76 (29)	
Hypertensive (Sys > 140 or Dia > 90)	78 (5)	48 (5)	30 (5)		12 (4)	18 (7)	
Booking Mean arterial pressure (mmHG)				0.527			0.587
Low (< 65)	26 (2)	16 (2)	10 (2)		4 (1)	6 (2)	
Normal (65–99)	1428 (93)	903 (93)	525 (92)		285 (93)	240 (91)	
High (> 100)	84 (5)	48 (5)	36 (6)		17 (6)	19 (7)	
Number of antenatal BP measures	5 (4–8)	5 (4–8)	6 (4–8)	0.008			
Any hypertension	337 (22)	208 (22)	129 (23)	0.620	62 (20)	67 (25)	0.311
Hypertension type[b] (%)							
Chronic	42 (13)	28 (13)	14 (11)	0.560	4 (7)	10 (15)	0.480
Gestational	284 (84)	172 (83)	112 (87)		56 (90)	56 (84)	
Unknown	11 (3)	8 (4)	3 (2)		2 (3)	1 (1)	
Diagnostic Criteria (%)				0.377			0.417
Both SBP ≥140 mmHg and DBP ≥90 mmHg	40 (12)	26 (13)	14 (11)		6 (10)	8 (12)	
Only SBP ≥140 mmHg	133 (39)	87 (42)	46 (36)		26 (42)	20 (30)	
Only DBP ≥90 mmHg	164 (49)	95 (46)	69 (53)		30 (48)	39 (58)	
Severity (%)				0.570			0.849
Non-severe: SBP 140–159 or DBP 90-109 mmHg	324 (96)	199 (96)	125 (97)		60 (97)	65 (97)	
Severe: BP ≥ 160 or DBP ≥110 mmHg)	253 (4)	9 (4)	4 (3)		2 (3)	2 (3)	

* Comparison between HIV-uninfected vs HIV-infected

** Comparison between HIV-uninfected vs HIV-infected (Initiated before pregnancy) vs HIV-infected (Initiated during pregnancy)

*** Best available measure (USS, SFH, LMP)

**** Chronic hypertension incidence estimated among women who were assessed at <20 weeks' gestation. Gestational hypertension incidence will be estimated among women previously assessed as normotensive at <20 weeks or assessed for the first time at 20 weeks.

6.4.1 Joint blood pressure trajectory groups

Preliminary analysis was conducted using single trajectory models, in order to identify the optimal degree of the polynomials; based on the lowest BIC criterion, the quadratic function fit the data best. For the joint SBP and DBP models using the lowest BIC criterion, the 5-class model was the best fitting (Suppl. Table 9.4.2), with model fit statistics indicating this to be an adequately fitting model (Suppl. Table 9.4.3). When assessing the probability matrix of membership for the five SBP and five DBP classes combinations, only seven of the possible 25 combinations had at least one individual with that joint trajectory (Suppl. Table 9.4.4). Based on clinical input, these seven SBP and DBP combinations were grouped and classified according to the longitudinal trend and the mean start BP. The four joint groups created were: consistent normal (SBP values persistently around 110mmHg, DBP values persistently around 70 mmHg; n=767 (44.3%)), consistent low normal (SBP values persistently around 100mmHg and DBP values persistently around 60mmHg; n=384 (23%)), consistent high normal (SBP values either around 110 mmHg or slightly increasing to between 120 -140 mmHg and DBP values either around 70 mmHg or slightly increasing to between 75-80 mmHg; n=316 (24%)), and increasing abnormal (both SBP and DBP values initially at normal levels then increasing (with differing magnitudes) during observation period; n=71 (9%)) BP (Figure 6.1).

Trajectory Groups	Group Membership
Consistent normal systolic and diastolic	44.3%
Systolic Group 3 (Green) + Diastolic Group 2 (Red) (44.3%)	
Consistent low normal systolic and diastolic	22.9%
Systolic Group 1 (blue) + Diastolic Group 3 (green) (22.9%)	
Consistent high normal systolic and diastolic	24.3%
Systolic Group 4 (yellow) + Diastolic Group 4 (yellow) (24.0%)	
Systolic Group 3 (green) + Diastolic Group 4 (yellow) (0.2%)	
Systolic Group 4 (yellow) + Diastolic Group 2 (Red) (0.1%)	
Increasing abnormal systolic and/or diastolic	8.5%
Systolic Group 5 (grey) + Diastolic Group 5 (grey) (6.1%)	
Systolic Group 4 (yellow) + Diastolic Group 1 (blue) (1.3%)	
Systolic Group 2 (red) + Diastolic Group 5 (grey) (1.1%)	

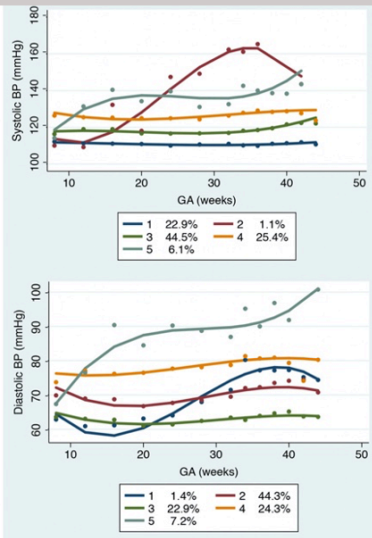

Figure 6.1 Joint trajectory group combinations

Five systolic (a) and five diastolic (b) blood pressure joint trajectory group combinations were grouped and classified according to longitudinal trend and mean start BP; into four joint groups: consistent normal systolic and diastolic BP, consistent low normal systolic and diastolic BP, consistent high normal systolic and diastolic BP, and increasing abnormal systolic and/or diastolic BP.

The abnormal trajectory group likely contains women who experience pre-eclampsia, which is characterised by a sudden, substantial increase in BP during the latter parts of pregnancy, along with evidence of end-organ and/or placental dysfunction (which we were unable to assess).

6.4.2 BP trajectory group predictors

Using the normal trajectory group as reference, women living with HIV were more likely to be in the low normal trajectory group (ARR 1.24, 95% CI 0.94-1.64) than women living without HIV (Suppl. Table 9.4.5). A higher proportion of older women (>30 years) were in the high normal and abnormal trajectory groups. Body mass index (BMI) status appeared to be associated with trajectory groups: women who were moderately and severely obese were more likely to be in the high normal (27% and 34% respectively) or abnormal (25% and 39% respectively) trajectory groups, compared to normal and overweight women who were more likely to be in the low normal (27% and 31% respectively) group (Suppl. Table 9.4.5). In multivariable analysis (adjusted for age, BMI, HIV status and gravidity), with the normal trajectory group as reference, older (ARR 1.52, 1.11-2.10) and obese (ARR 2.06, 1.31-3.25) women were at increased risk of being in the high normal group.

Among women living with HIV, no differences were observed according to timing of initiation in the normal, low normal and high normal trajectory groups. However among the women with abnormal trajectories, they were more likely to have initiating ART during pregnancy compared to initiating preconception (20% vs 10%) (Suppl. Table 9.4.6).

About a third of women in the high normal (35%) and abnormal trajectory groups (31%) were classified as being normotensive at their first ANC, highlighting the problem with only using cross-sectional BP assessments. A significant proportion of women in the high normal and abnormal trajectory groups were classified as being pre-hypertensive at the first ANC, 53% and 42% respectively (Suppl. Table 9.4.6). The majority of women classified as being hypertensive at the first ANC were in the high normal (11%) and abnormal (27%) trajectory groups; with a third of women in abnormal trajectory group also classified has having a high mean arterial pressure.

When hypertension was assessed throughout the antenatal period, the majority of women in the high normal (51%) and abnormal (97%) trajectory group met the conventional hypertensive cut off at least once (Table 6.2). When grouped according to onset of hypertension the majority of women with hypertension in all trajectory groups were classified as having gestational hypertension.

Table 6.2 Hypertension description by trajectory group

	Total *N* = 1538	Group 1: Normal *N* = 767	Group 2: Low Normal *N* = 384	Group 3: High Normal *N* = 316	Group 4: Abnormal *N* = 71	*P* value
Number of measures	5 (4–8)	5 (4–7)	5 (4–7)	6 (4–10)	7 (4–16)	< 0.0001
Any hypertension (%)	337 (22)	93 (12)	14 (4)	161 (51)	69 (97)	< 0.0001
Hypertension type* (%)						
Chronic	42 (13)	15 (16)	2 (14)	23 (14)	2 (3)	0.001
Gestational	284 (84)	70 (75)	11 (79)	136 (85)	67 (97)	
Unknown	11 (3)	8 (9)	3 (2)	2 (1)	0	
Diagnostic Criteria (%)						< 0.0001
Both SBP ≥140 mmHg and DBP ≥90 mmHg	40 (12)	5 (5)	0	12 (7)	23 (33)	
Only SBP ≥140 mmHg	133 (39)	46 (50)	7 (50)	66 (41)	14 (20)	
Only DBP ≥90 mmHg	164 (49)	42 (45)	7 (50)	83 (52)	32 (47)	
Severity (%)						0.578
Non-severe (SBP 140–159 or DBP 90-109 mmHg)	324 (96)	89 (96)	13 (93)	157 (98)	65 (94)	
Severe (SBP ≥ 160 or DBP ≥110 mmHg)	253 (4)	4 (4)	1 (7)	4 (2)	4 (6)	

* Chronic hypertension incidence estimated among women who were assessed at <20 weeks' gestation. Gestational hypertension incidence will be estimated among women previously assessed as normotensive at <20 weeks. Unknown = assessed for the first time at 20 weeks.

6.4.3 BP Trajectory groups and adverse birth outcomes

Comparing outcomes overall between the joint trajectory groups, a higher incidence of any PTD (unadjusted, RR3.13, 95% CI 2.35-4.17; 50%) and any LBW (RR 2.81, 95% CI 1.90-4.15; 32%) was observed among women in the abnormal trajectory group. The lowest risk of PTD was observed in the consistent low normal group (RR 0.67, 95% CI 0.48-0.94; 11%) (Table 6.3; Table 6.4). There was a fairly even split between late and moderately preterm among the normal (37% vs 43%) and low normal (46% vs 34%); trajectory groups, but in contrast there was a higher incidence of moderate than late preterm in the high normal (53% vs 25%) and abnormal (46% vs 26%) trajectory groups. Further, 29% of the preterm deliveries in the abnormal trajectory group were classified as very preterm (Table 6.3). Among women with PTD, there was no appreciable difference in delivery mode of women in the normal and high normal trajectory groups; however in contrast in the low normal and abnormal trajectory groups PTD was commonly associated with emergency caesarean section, indicating a suspicion of adverse pregnancy outcome necessitating preterm delivery.

Table 6.3 Birth Outcomes by joint trajectory groups

	Total	Group 1: Normal	Group 2: Low Normal	Group 3: High Normal	Group 4: Abnormal	*P* value*
Gestational Age (weeks) (%)	*N* = 1532	*N* = 764	*N* = 382	*N* = 316	*N* = 70	< 0.0001
Term (≥ 37)	1262 (82)	642 (84)	341 (89)	244 (77)	35 (50)	
Any Preterm (< 37)	270 (18)	122 (16)	41 (11)	72 (23)	35 (50)	
Late Preterm (34–37)	93 (6)	45 (6)	19 (5)	20 (7)	9 (13)	
Moderately Preterm (32–34)	120 (8)	52 (7)	14 (4)	38 (12)	16 (23)	
Very Preterm (28–32)	57 (4)	25 (3)	8 (2)	14 (4)	10 (14)	
Birthweight (g) (%)	*N* = 1530	*N* = 763	*N* = 384	*N* = 312	*N* = 71	< 0.0001
Normal (≥2500)	1333 (87)	675 (88)	342 (89)	268 (86)	48 (68)	
Any LBW (< 2500)	197 (13)	88 (12)	42 (11)	44 (14)	23 (32)	
LBW (< 2500)	159 (10)	75 (10)	37 (10)	34 (11)	13 (18)	
Very LBW (< 1500)	38 (3)	13 (2)	5 (1)	10 (3)	10 (14)	
Mean (SD)	3116 (644)	3146 (616)	3158 (550)	3110 (686)	2580 (934)	

All the variables had <4% missing data, with similar proportions of missing data across the comparison groups

In the normal, low normal and high normal trajectory groups the incidence of LBW was similar (12% vs 11% vs 14%), whereas a significantly higher incidence was observed in the abnormal trajectory group (32%). In the normal, low normal and high normal trajectory groups, of the LBW neonates, 12% and 23% were VLBW, whereas in the abnormal trajectory group 44% were VLBW (Table 6.3). In multivariable analyses allowing for other variables potentially associated with PTD or LBW, with the normal group as reference, a low normal trajectory (ARR0.59, 0.41-0.85) was associated with decreased risk of PTD, high normal (ARR1.48, 1.12-1.95) and abnormal trajectories with increased risk of PTD (ARR3.18, 2.32-4.37). The abnormal trajectory was also significantly associated with increased LBW risk (ARR1.78, 1.22-2.59) (Table 6.4).

Table 6.4 Adjusted Associations between joint trajectory groups and adverse birth outcomes

Trajectory Group (Ref cat: Normal)	Preterm[a]			Low Birthweight[b]		
	RR (95% CI)	ARR (95% CI)	*P*-value	RR (95% CI)	ARR (95% CI)	*P*-value
Low Normal	0.67 (0.48–0.94)	0.59 (0.41–0.85)	0.004	0.95 (0.67–1.34)	0.99 (0.70–1.43)	0.998
High Normal	1.43 (1.10–1.85)	1.48 (1.12–1.95)	0.005	1.22 (0.87–1.71)	1.16 (0.83–1.61)	0.390
Abnormal	3.13 (2.35–4.17)	3.18 (2.32–4.37)	< 0.0001	2.81 (1.90–4.15)	1.78 (1.22–2.59)	0.003

* adjusted for age, body mass index and previous preterm delivery
** adjusted for age, body mass index and gestational age at delivery

In stratified analyses to explore potential effect modification by HIV status, with the normal group as reference, among women living without HIV low normal trajectory was associated with significantly decreased risk of PTD (ARR0.48, 0.29-0.81) while high normal (ARR1.53, 1.08-2.16) and abnormal trajectories (ARR3.42, 2.31-5.05) were associated with significantly increased risk of PTD. Among women living with HIV these associations were similar but did not reach statistical significance (low normal and PTD: ARR 0.74, 95% CI 0.44-1.26; high normal and PTD: ARR1.40 95% CI 0.90- 2.19; abnormal and PTD: ARR2.66 95% CI 1.49-4.76) (Table 6.5). The limited sample size of women living with HIV could have contributed to the imprecision of the stratified estimates.

There appeared to be evidence of effect measure modification on the relative risk scale by HIV status in the relationship between the abnormal trajectory group and LBW infants: among women living without HIV in the abnormal trajectory group there was an increased risk of LBW (ARR2.01, 95% CI 1.33-3.02), while in the women living with HIV this effect was attenuated and did not reach statistical significance (ARR1.08, 95% CI 0.43-2.72) (Table 6.5).

Table 6.5 Adjusted Associations between joint trajectory groups and adverse birth outcomes stratified by HIV status

Trajectory Group (Ref cat: Normal)	Preterm[a]			Low Birthweight[b]		
	HIV-uninfected ARR (95% CI)[†]	HIV-infected ARR (95% CI)[††]	*P*-value	HIV-uninfected ARR (95% CI)[†]	HIV-infected ARR (95% CI)[††]	*P*-value
Low Normal	0.48 (0.29–0.81)	0.74 (0.44–1.26)	0.005[†] 0.267[††]	0.75 (0.46–1.22)	1.37 (0.80–2.33)	0.244[†] 0.252[††]
High Normal	1.53 (1.08–2.16)	1.40 (0.90–2.19)	0.017[†] 0.134[††]	1.13 (0.75–1.70)	1.18 (0.70–2.01)	0.549[†] 0.531[††]
Abnormal	3.42 (2.31–5.05)	2.66 (1.49–4.76)	< 0.0001† 0.001[††]	2.01 (1.33–3.02)	1.08 (0.43–2.72)	0.001[†] 0.872[††]

* adjusted for age, body mass index and previous preterm delivery
** adjusted for age, body mass index and gestational age at delivery
[†] p-value for the association between adverse birth outcome and trajectory group categories among HIV-uninfected women
[††] p-value for the association between adverse birth outcome and trajectory group categories among HIV-infected women

6.5 Discussion

In this prospective cohort of pregnant women living with and without HIV at a large public sector primary care antenatal facility in South Africa, we identified groups with distinct antenatal BP trajectories, which were not associated with HIV status; both high consistent and increasing abnormal BP trajectories were associated with subsequent risk of PTD and LBW. Understanding different BP trajectories in pregnancy, especially in areas with high HIV prevalence, which has been shown to be associated with BP, will inform efforts to identify women at increased risk of developing HDP and adverse birth outcomes.

Research studies tend to use stringent criteria (traditional cut-offs), with a strong emphasis on correctly ruling out those without hypertension from being classified as hypertensive. Additionally, these studies usually predict subsequent HDP development based on BP measured at a single time point such at the first ANC visit (23, 24). Using such a cross-sectional approach, 5% of women in our cohort met the hypertension cut-off at their first ANC visit while 22% were classified as hypertensive at any point in pregnancy; substantially higher than levels observed in a trial conducted in southern Asia and sub-Saharan Africa (25). In contrast, clinically all-encompassing and practical criteria are predominantly used to classify HDP, even if they have higher false positivity rates, because the ultimate goal is to identify women with a higher-than-average risk to themselves or their fetuses in order to guide clinical management (26). This approach is more informative as it differentiates between new onset HDP, which can affect maternal and infant morbidity and mortality (27), and isolated new-onset SBP and/or DBP elevation, which is quite common during pregnancy (28).

The GBTM approach is well-suited to explore trajectories for outcomes that are commonly categorised, such the natural history of BP in pregnancy. Using this approach, we identified distinct trajectory groups which we classified according to the mean BP values relative to clinical definitions of BP levels and the pattern of change throughout the antenatal period (consistent or increasing). The patterns observed in the low normal, normal and high normal trajectory groups resembles patterns seen in clinically healthy women, where BP decreases steadily until mid-pregnancy then rises again returning to preconception levels in late pregnancy until delivery (30). In contrast, clinical reports show that in women who develop gestational hypertension, BP remains stable until mid-pregnancy then increases until delivery (29, 30) – the abnormal trajectory group resembles this pattern. This group appeared to be a combination of women with chronic hypertension (likely classified as hypertensive at baseline (27%)), and women with gestational hypertension (42% prehypertensive, 31% normotensive at baseline). These findings suggest that similar to the approach applied clinically, a better approach for risk assessment in research studies would be incorporation of BP measurements throughout pregnancy to assess trajectory of BP before delivery rather than the traditional cross-sectional measures.

We found that BP characteristics were similar in women living with and without HIV, half of whom were on ART from before pregnancy and half from ~16 weeks of pregnancy, with no differences in the incidence of hypertension at baseline or throughout pregnancy. Studies investigating the association between HIV and HDP have shown conflicting results with either increased risk (7), or no

difference in risk (31, 32). Though limited, data on specific ART drug classes has shown exposure to PIs to be strongly associated with hypertension (7, 33); in our setting nucleoside reverse transcriptase inhibitors regimens are predominantly used rather than PIs which could explain lack of difference by HIV status. Our findings by HIV status are similar to a retrospective cohort study of two overlapping pregnancy cohorts (34) but differed according to ART status. We found that women classified as hypertensive or as having abnormal trajectories were more likely to have initiated ART in pregnancy; and this could possibly be linked to immune reconstitution, a hyper-inflammatory state associated with ART initiation (35, 36), vascular dysfunction and alterations in sympathetic nervous outflow (37). This process is exacerbated by the presence of well-known risk factors including extreme reproductive ages, obesity and genetic predisposition. In our cohort, we confirm increased maternal age and BMI to be associated with hypertension, consistent with other studies (38, 39).

High blood pressure in pregnancy is associated with adverse pregnancy outcomes for both mother and infant. Similar to findings from Ethiopia, we found that approximately a third of women classified as hypertensive at 1st ANC visit delivered preterm and LBW infants (40).
However, given that only 5% of women were classified as hypertensive at first ANC (using traditional cut-offs) this cross-sectional assessment approach may not correctly identify women at increased adverse birth outcome risk. In contrast, using GBTM enabled identification of women who were not classified as hypertensive but who were at increased risk, as such the HDP and adverse birth outcomes association is likely to be more accurate using this approach.
We found that women with high normal and abnormal trajectories were at increased risk of PTD, while those with abnormal trajectories were at increased LBW risk. This is consistent with previous data showing that among women without pre-existing hypertension or preeclampsia, higher maternal blood pressure levels were associated with impaired fetal growth from the third trimester onward as well as increased risks of adverse birth outcomes (29, 41). Increased PTD incidence among women with abnormal BP levels is expected because of pre-emptive clinical management protocols that prioritise maternal safety, and thus may necessitate preterm delivery to prevent severe complications of HDP, which could explain the association between PTD and emergency caesarean sections in our cohort. The association between LBW and HDP has been postulated to be attributable to both the effect of early delivery and/or fetal growth restriction (42). The mechanisms by which maternal blood pressure levels can affect fetal growth are unclear, with suggestions that it could be related to placental dysfunction or adverse maternal cardiovascular adaptations to pregnancy (29). Separately, associations between HIV/ART and BP and adverse birth outcomes have

been observed, however in our study where we assessed HIV status as an effect modifier of the relationship between BP trajectories and adverse birth outcomes, we observed no differences by HIV status. While associations were attenuated for women living with HIV, we hypothesise that limited sample size rather than an HIV effect are the explanation - an assumption further substantiated by the lack of BP differences by HIV status.

This analysis is among the first to explore discrete latent class assignment of BP trajectory groups during pregnancy and their relationship with adverse birth outcomes, in a population-based study. Strengths of the study include the use of prospectively collected blood pressure and gestational age measurements at baseline and throughout pregnancy, which enabled the distinction between women with chronic and gestational hypertension based on gestation-dependent BP changes. Additionally, GBTM is an advantageous analytic approach because it does not rely on predefined categorisation of trajectories, removing subjectivity regarding class assignment, and enabling capturing of within-class heterogeneity that might be missed with analyses based on class-averaged means (20). It should however be noted that the trajectory grouping using this approach depends on the data, and while the trajectory groups may provide a general trend of BP control over time, they may not accurately characterise each individual's actual BP trajectory.

A limitation of this analytical approach is that it does not enable interpretation of blood pressure in pregnancy by clinicians in the light of international gestational age-specific centiles. An alternate approach to be used in future studies determining gestational hypertensive disorders would involve defining centiles for systolic and diastolic blood pressure across healthy pregnancies. Additional limitations include the availability of only basic maternal characteristics for use in adjusted analyses with no data on other pre-pregnancy risk factors for HDP such as hypertension family history, previous HDP in multipara women; and no proteinuria data to establish whether the women in abnormal trajectory group had preeclampsia. Lacking detailed preconception and early pregnancy maternal data, we were unable to find significant association with baseline characteristics and trajectory, except for maternal BMI and gravidity. Additionally, blood pressure measurements were collected during routine clinical practice, where the use of a single BP measure may have greater measurement error than standardised measures. This analysis provides some of the first insights into the relationships between BP trajectory levels and their association with adverse birth outcomes, in particular the use of trajectories enables identification of women at increased risk of adverse birth outcomes who would not be picked up using the traditional approach currently employed in studies investigating this association.

6.6 Conclusion

Globally, HDP incidence has increased over time and as obesity increases in low- and middle-income countries, high blood pressure during pregnancy has the potential to adversely impact maternal and child health outcomes. It is important for public health to monitor trends in pregnancy outcomes by HDP. We detected heterogeneity in general relationship between HDP and PTD/LBW outcomes, by identifying five groups with distinctive patterns of BP trends over the pregnancy. Further methodology work is required with larger datasets in order establish where precisely divergence between trajectories occurs in order to inform optimum timing of BP measurements in studies investigating HDP. Larger datasets would also enable identification of heterogeneity within trajectory groups particularly among the high normal and abnormal trajectories. Further work is also required to inform understanding of different BP trajectories in pregnancy, particularly in high HIV prevalent settings where ART use is linked to weight gain which may extenuate BP levels, in order to provide appropriate surveillance and counselling to pregnant women living with HIV and those with pregnancy intentions.

6.7 References

1. Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy. *Am J Obstet Gynecol*. 2000;183(1):s1-s22.
2. Magee LA, Abalos E, von Dadelszen P, et al. How to manage hypertension in pregnancy effectively. *Br J Clin Pharmacol.* 2011;72(3):394-401.
3. Brown MA, Lindheimer MD, de Swiet M, Assche AV, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). Taylor & Francis; 2001.
4. Koopmans CM, Bijlenga D, Groen H, et al. Induction of labour versus expectant monitoring for gestational hypertension or mild pre-eclampsia after 36 weeks' gestation (HYPITAT): a multicentre, open-label randomised controlled trial. *Lancet.* 2009;374(9694):979-988.
5. Browne J, Vissers K, Antwi E, et al. Perinatal outcomes after hypertensive disorders in pregnancy in a low resource setting. *Trop Med Int Health*. 2015;20(12):1778-1786.
6. Paller MS. Hypertension in pregnancy. *J Am Soc Nephrol*. 1998;9(2):314-321.
7. Premkumar A, Dude AM, Haddad LB, Yee LM. Combined antiretroviral therapy for HIV and the risk of hypertensive disorders of pregnancy: A systematic review. *Pregnancy Hypertens*. 2019;17:178-190.
8. Brocklehurst P, and French, R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *Br J Obstet Gynaecol*. 1998;105(8):836-848.
9. Rollins NC, Coovadia HM, Bland RM, et al. Pregnancy outcomes in HIV-infected and uninfected women in rural and urban South Africa. *J Acquir Immune Defic Syndr.* 2007;44(3):321-328.
10. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.
11. Li N, Sando MM, Spiegelman D, et al. Antiretroviral therapy in relation to birth outcomes among HIV-infected women: a cohort study. *J Infect Dis.* 2016;213(7):1057-1064.
12. Malaba TR, Phillips TK, Le Roux S, et al. Antiretroviral therapy use during pregnancy and adverse birth outcomes in South African women. *Int J Epidemiol*. 2017;46(5):1678-1689.
13. Nathan HL, Duhig K, Hezelgrave NL, Chappell LC, Shennan AH. Blood pressure measurement in pregnancy. *Obstet Gynaecol.* 2015;17(2):91-98.
14. Maddox TM, Ross C, Tavel HM, et al. Blood pressure trajectories and associations with treatment intensification, medication adherence, and outcomes among newly diagnosed coronary artery disease patients. *Circ Cardiovasc Qual Outcomes*. 2010;3(4):347-357.
15. Malaba TR, Gray CM, Myer L, Newell ML. Cohort Profile: Prematurity Immunology in HIV-infected Mothers and their infants Study (PIMS). *BMJ Open*. 2021;11(9):e047133

16. Chobanian AV, Bakris GL, Black HR, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560-2571.
17. Brown MA, Magee LA, Kenny LC, et al. Hypertensive disorders of pregnancy: ISSHP classification, diagnosis, and management recommendations for international practice. *Hypertension.* 2018;72(1):24-43.
18. DeMers D, Wachs D. Physiology, Mean Arterial Pressure. In: StatPearls [Internet]. Florida, USA: StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK538226/ (accessed 18 October 2019)
19. L Jones B, Nagin D. A Stata plugin for estimating group-based trajectory models. 2012.
20. Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol Methods*. 2001;6(1):18-34.
21. Helgeson VS, Snyder P, Seltman H. Psychological and physical adjustment to breast cancer over 4 years: identifying distinct trajectories of change. *Health Psychol*. 2004;23(1):3-15.
22. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* 2004;159(7):702-706.
23. North RA, McCowan LM, Dekker GA, et al. Clinical risk prediction for pre-eclampsia in nulliparous women: development of model in international prospective cohort. *BMJ*. 2011;342:d1875.
24. Broughton PF, Sharif J, Lal S. Predicting high blood pressure in pregnancy: a multivariate approach. *J Hypertens*. 1998;16(2):221-229.
25. Magee LA, Sharma S, Nathan HL, et al. The incidence of pregnancy hypertension in India, Pakistan, Mozambique, and Nigeria: A prospective population-level analysis. *PLoS Med*. 2019;16(4):e1002783.
26. Higgins JR, de Swiet M. Blood-pressure measurement and classification in pregnancy. *Lancet.* 2001;357(9250):131-135.
27. Hutcheon JA, Lisonkova S, Joseph K. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol.* 2011;25(4):391-403.
28. Regitz-Zagrosek V, Blomstrom Lundqvist C, Borghi C, et al. European Society of Gynecology (ESG), Association for European Paediatric Cardiology (AEPC), German Society for Gender Medicine (DGesGM). ESC Guidelines on the management of cardiovascular diseases during pregnancy: the Task Force on the Management of Cardiovascular Diseases during Pregnancy of the European Society of Cardiology (ESC). *Eur Heart J.* 2011;32(24):3147-3197.

29. Bakker R, Steegers EA, Hofman A, Jaddoe VW. Blood pressure in different gestational trimesters, fetal growth, and the risk of adverse birth outcomes: the generation R study. *Am J Epidemiol*. 2011;174(7):797-806.
30. Hermida RC, Ayala DE, Iglesias M. Predictable blood pressure variability in healthy and complicated pregnancies. *Hypertension.* 2001;38(3):736-741.
31. Bodkin C, Klopper H, Langley G. A comparison of HIV positive and negative pregnant women at a public sector hospital in South Africa. *J Clin Nurs*. 2006;15(6):735-741.
32. Shapiro RL, Souda S, Parekh N, et al. High prevalence of hypertension and placental insufficiency, but no in utero HIV transmission, among women on HAART with stillbirths in Botswana. *PLoS One*. 2012;7(2): e31580.
33. Tripathi A, Jerrell JM, Skelton TN, Nickels MA, Duffus WA. Incidence of primary hypertension in a population–based cohort of HIV–infected compared with non–HIV–infected persons and the effect of combined antiretroviral therapy. *J Am Soc Hypertens*. 2015;9(5):351-357.
34. Saums MK, King CC, Adams JC, et al. Combination Antiretroviral Therapy and Hypertensive Disorders of Pregnancy. *Obstet Gynecol*. 2019;134(6):1205-1214.
35. Wimalasundera R, Larbalestier N, Smith J, et al. Pre-eclampsia, antiretroviral therapy, and immune reconstitution. *Lancet.* 2002;360(9340):1152-1154.
36. Sebitloane HM, Moodley D. The impact of highly active antiretroviral therapy on obstetric conditions: a review. *Eur J Obstet Gynecol Reprod Biol.* 2017;210:126-131.
37. Masenga SK, Hamooya BM, Nzala S, et al. Patho-immune mechanisms of hypertension in HIV: A systematic and thematic review. *Curr Hypertens Rep.* 2019;21(7):56-75.
38. Kahsay HB, Gashe FE, Ayele WM. Risk factors for hypertensive disorders of pregnancy among mothers in Tigray region, Ethiopia: matched case-control study. *BMC Pregnancy Childbirth*. 2018;18(1):482-492.
39. Lee SK, Kim SH, Cho G-Y, et al. Obesity phenotype and incident hypertension: a prospective community-based cohort study. *J Hypertens*. 2013;31(1):145-151.
40. Berhe AK, Ilesanmi AO, Aimakhu CO, Mulugeta A. Effect of pregnancy induced hypertension on adverse perinatal outcomes in Tigray regional state, Ethiopia: a prospective cohort study. *BMC Pregnancy Childbirth*. 2020;20(7):1-11.
41. Macdonald-Wallis C, Silverwood RJ, Fraser A, et al. Gestational-age-specific reference ranges for blood pressure in pregnancy: findings from a prospective cohort. *J Hypertens*. 2015;33(1):96-105.
42. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull WHO*. 1987;65:663–737.

Chapter 7: Impact of data sources and selection bias on the association between maternal ART use and preterm delivery

Malaba TR, Kalk E, Madlala HP, Davies MA, Boulle A, Myer L. Impact of data sources and selection bias on the association between maternal ART use and preterm delivery. *Manuscript to be submitted in 2022.*

Relevance of this paper to the thesis

This chapter presents data on the impact of different data sources on the reported incidence of preterm delivery. There is exploration of the potential impact of bias related to participant selection and data quality from different data sources on the association between maternal HIV and antiretroviral therapy use and preterm delivery. While bias has been hypothesised as a methodological issue that could explain the heterogeneity in findings in previous studies, bias related to the factors investigated in this chapter have not previously been explored in relation to this association.

Contribution of the student and co-authors

TRM conceptualised and conducted all analyses, led data interpretation and drafted the manuscript. HM directed data collection, with assistance from TRM. AB, LM and MAD conceptualised the study from which these data arise, were responsible for funding, implementation and overall leadership, and assisted with data interpretation. All authors have contributed to this draft manuscript.

7.1 Abstract

7.1.1 Background

The majority of studies investigating the association between maternal antiretroviral therapy (ART) use and preterm delivery (PTD) tend to be observational and thus subject to bias. This is thought to be one factor contributing to the heterogeneity, with bias related to participant selection and data quality from different data sources not explored in relation to this association.

7.1.2 Methods

Pregnant women at public sector antenatal facilities in the Western Cape Province were enrolled into a multi-component study, which collected data from the same underlying population over the same time period. The three related data sources were based on a province-wide pregnancy surveillance, a facility-based pregnancy exposure registry (PER) and a prospective cohort study. Each study component had different levels of data availability and quality depending on the data sources. Surveillance data was based on routine electronic health records which were linked across health services by the Provincial Health Data Centre (PHDC). Data for the PER data was based on the digitalisation of the patient held booklet containing medical and obstetric history, and all pregnancy-related clinical consultations and investigations throughout pregnancy. The cohort study data was based on active follow up of women throughout pregnancy with questionnaire-based data collection as well as data abstraction of clinical records. The impact of the different data sources on the associations between maternal ART use with preterm delivery was examined separately using modified Poisson regression with robust error variance.

7.1.3 Results

The three data sources included 183 593 women in the PHDC population (19% living with HIV), 9476 in the PER population (32% living with HIV) and 989 in the prospective cohort study (48% living with HIV). Among women with live singleton births, adjusted risk ratio (ARR) of preterm delivery was 1.15 (95% Confidence Interval (CI) 1.11 - 1.18) among women living with HIV compared to their counterparts living without HIV in the PHDC population. The associations observed in the PER and cohort study populations were less pronounced, with no significant differences seen by HIV status and timing of ART initiation. Differences in the representativeness of women living with HIV seen across the data sources and differing data quality did not appear to impact the association estimates observed.

7.1.4 Conclusion

In this comparison of routine care cohorts across different data sources, the risk of PTD was increased in women living with HIV in the larger PHDC population. These findings demonstrate that in future investigations of the association between HIV/ART status and adverse birth outcomes it will be important to include both population level data and data from follow-up studies in order to leverage the benefits of both sets of data sources.

7.2 Introduction

Inconsistent findings, across different settings, in the association between maternal antiretroviral therapy (ART) use and adverse birth outcomes, especially preterm delivery (PTD), are at least partially explained by study designs used. Pregnant women have generally been excluded from clinical trials, so data investigating possible ART-associated adverse birth outcomes comes primarily from observational studies, which can be subject to substantial bias and confounding (1, 2). Selection bias is one such bias which, like confounding, generates pathways between the exposure and outcome that are non-causal, and this distortion can result in the masking of a true association or a true lack of association (3). Any such bias is a major concern in studies where the aim is to provide the information that informs the identification of target groups in order to guide clinical care or inform understanding of why there is an association.

To understand the predictors and antecedents of any adverse birth outcomes and related biological processes, it would be advantageous to study these in representative samples of pregnancies (4). However, unrepresentative study populations are common in perinatal epidemiology, because of the high natural attrition that occurs from the preconception period through to birth. Consequently underlying populations are incompletely observed complicating the identification of the study population (5). The process of recruiting women into studies investigating maternal ART use and adverse birth outcomes can induce bias, because of natural attrition as well as the propensity to enrol women at convenient time points such as during visits to antenatal care services. If the exposure, outcome, and/or their correlates, influence study eligibility or study participation, the resulting exposure–outcome association may not be a true reflection of the causal effect in the underlying population (6). Selection bias can also manifest when population subgroups have varying participation rates and differing exposure-outcome associations (7). For example, pregnant women who do not access any health care during pregnancy including at delivery, as well as those who only access these facilities at delivery, are typically excluded from study populations, which could also lead to unrepresentative study samples.

While observational studies enable inclusion of populations in a “real-world” context, the presence of selection bias, unmeasured confounding, and inappropriate effect definition and estimation could hamper the ability to make reliable inferences regarding causality from their findings (8). In recent years, in an effort to circumvent the limitations of smaller observational studies and simultaneously improve data quality through larger datasets, the safety of medications in pregnancy such as ART has increasingly been assessed on a population level using electronic health data from large national

pregnancy medication exposure registries (9, 10) and administrative databases (11-13). Accurate determination of the contribution of an exposure to an outcome depends on the robust measurement of preconception and/or pregnancy exposures, outcomes of interest and relevant covariates. These new population level data sources, through inclusion of people who are not typically enrolled in epidemiological studies, have improved the data quality, albeit on a limited set of variables, and are valuable for public health, where the focus is picking up on patterns and trends rather than establishing causal links. It should however be noted that this routinely collected data is not intended for research, and it is conceivable that bias (selection and information), one of the many key challenges to making valid inferences, is becoming increasingly prevalent in the study of birth outcomes using these data sources. This would have a profound effect on the results and their interpretation because bias impacts both the internal and external validity (14).

Since not all sources of bias in studies threaten the validity of findings, it is essential to investigate the potential impact of biases in relation to effect estimates (8). In a recent study, where we accounted for gestational age measurement error, we found no association between maternal ART (predominantly with a regimen of tenofovir (TDF), emtricitabine (FTC) and efavirenz (EFV)) and preterm delivery (15). It is not clear, however, whether these results are a reflection of the truth or whether it could be due to limited sample size and/or bias impacting effect estimates with movement towards the null. The existence of selection bias and any impact it might have can at times be more difficult to recognise than known confounding or measurement error (16), because, by definition, data on those not included are unavailable so selection bias can typically only be hypothesised (17, 18). One method that has been used to investigate selection bias is through the comparison of exposure-outcome associations from study participants to similar associations obtained from the underlying population using registry data or administrative databases available for both populations (18-21). Since these data are collected independently with no regard for potential participation in future epidemiological studies, any presence of bias in effect estimates likely reflects selection bias. Additionally, since cohort studies are usually intended to reflect the underlying population, this approach could also serve as a direct measure of generalisability of the study results (18).

The Western Cape Provincial Health Data Centre (PHDC) is a health information exchange leveraged on the use of a unique patient identifier allocated to users of public sector health services. Patient records are linked across health services which provides a valuable opportunity to investigate whether cohort formation induces any selection bias. Using different data sources in the same

population over the same time period, with different availability and quality of exposures, outcome and confounding variables; we investigated the association between maternal ART on PTD in Cape Town, South Africa.

7.3 Methods

7.3.1 Data sources

Data for these analyses were drawn from B-Positive, apopulation-level study, which sought to explore the population-level impact of the universal initiation of lifelong ART policy for pregnant women living with HIV while strengthening routine data systems (Figure 7.1). B-Positive was a multi-component study, with data from population of women included coming from different but related sources (provincial and sentinel site level) over the same time period. Pregnant women (>18 years) accessing health care at fixed public sector services and facilities (primary care clinics, secondary and tertiary level hospitals) in the Western Cape Province were enrolled into the various components of the study between January 2017 and July 2018.

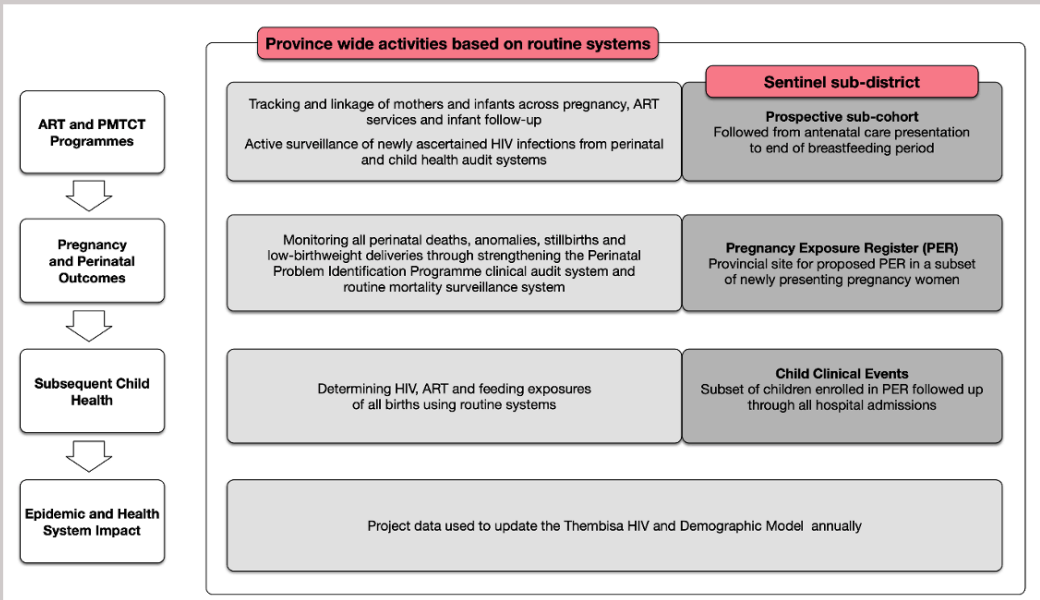

Figure 7.1 Outline of B-Positive study activities at both provincial and sentinel site (Gugulethu Midwife Obstetric Unit) level

ART – antiretroviral therapy, PMTCT – prevention of mother to child transmission

One study component of the B-Positive study was the province-wide surveillance of the pregnancy and perinatal outcomes of pregnant women (data source 1). This was facilitated by the Western Cape PHDC, a patient-level interlinked health information exchange leveraged on the WC's unique patient identifier allocated to all patients accessing public sector health services to facilitate linkage of individual patients across electronic record systems (22). Using this identifier, various routine electronic clinical, clerical and administrative database platforms across the province were centrally

combined to generate an individual-level population database which was used in this analysis (PHDC population) (22, 23). An algorithm was used by the PHDC to identify pregnancy events based on antenatal clinic visit attendance, presence of rhesus antibody test, pregnancy or abortive outcome of pregnancy indicated by an International Classification of Diseases code, pregnancy termination drugs dispensed, or recording of a birth on the facility birth register (24). Based on this information, pregnancy outcomes and the corresponding outcome dates were inferred, where possible. To distinguish between a missing outcome and a non-pregnancy event, a pregnancy confidence score was calculated in order to measure the degree of confidence that a pregnancy had truly occurred (24).

The other two components of the B-Positive study were dedicated sentinel studies, which were linked and paired with province-wide surveillance activities (Figure 7.1). These study activities were conducted in a single geographic area to investigate population-level trends in greater detail. The sentinel surveillance activities were conducted at the Gugulethu Midwife's Obstetric Unit (GMOU), a large public sector primary care facility that provides antenatal care and serves a peri-urban community with predominantly low-income, high HIV prevalence sub-district of Cape Town, South Africa. These activities include the establishment of a pregnancy exposure registry (PER) and a prospective cohort study, which both enrolled pregnant women accessing antenatal care (ANC) at the GMOU. The PER, which served as a proposed provincial site for a national pregnancy registry and birth defects surveillance, digitised data elements from the maternity case record (MCR). This is a patient-held booklet that records medical and obstetric history, and all pregnancy-related clinical consultations and investigations at the first ANC visit, updated at all subsequent ANC visits (25). Women were entered into the PER at their first visit to GMOU, with data collected using the electronic primary care information system that is present in all WC public health facilities (PER population) (data source 2). The prospective cohort study conducted at GMOU enrolled consecutive pregnant women regardless of HIV status attending their first antenatal visit at the facility. Participants in the cohort study were actively followed up through up to three intensive antenatal study visits, depending on the gestational age at enrolment. In addition to data abstraction from antenatal, obstetric, medical and laboratory records, data was also collected through questionnaires, administered by trained study interviewers, which included maternal demographics, pregnancy history and healthcare information (data source 3). Data collected through the questionnaires provided more detailed and nuanced information.

7.3.2 Exposures and Outcomes

The exposures of interest were HIV status, categorised as living with or without HIV , and ART status categorised as ART initiation before or during pregnancy. Exposure status categorisation the PHDC population was based a combination of the presence of a record of a positive enzyme-linked immunosorbent assay (ELISA), a $CD4^+$ T-cell count, a viral load test result, dispensing of combination ART, and/or registration on databases for the management of individuals living with HIV (24). The date of ART initiation was determined by taking the earlier of the first recorded date on dispensing of ART as date of ART initiation (24). For the PER and cohort study populations, exposure status was based on assessments conducted and recorded in the MCR during the first ANC visit, including HIV testing for women without a previous HIV diagnosis. In the cohort study, study questionnaires were able to determine which ART regimens the women were using. During the study period, women newly diagnosed with HIV infection and those previously diagnosed but not on ART were initiated on a fixed-dose combination of TDF+FTC+EFV at the first ANC visit (ART initiation during pregnancy). Women who conceived while on ART continued their current regimen throughout pregnancy (ART initiation before pregnancy).

The primary outcome of interest was gestational age at delivery, with birthweight a secondary outcome of interest. Among all live singleton births, gestational age was defined as completed weeks at delivery, categorised as a preterm (<37), term (≥37-<42) or post-term (≥42) delivery (26). Preterm delivery was further subdivided into late/moderately preterm (32–<37), very preterm (28-<32), extremely preterm (<28) delivery. As part of routine care services, all pregnant women have their gestational age clinically assessed by a midwife through last menstrual period (LMP) and/or measurement of symphysis fundal height (SFH); in facilities with ultrasound a small subset of women may be referred for an ultrasound assessment when estimated dates of delivery between LMP and SFH do not match. For the PHDC population, gestational age at delivery was calculated based on the best available measures. In cases where there were concerns around the quality of the gestational age data, a back calculation was used based on birthweight before the data was obtained. For the PER population, gestational age at delivery was calculated based on routine antenatal care ultrasound or the combination of LMP/SFH, particularly at later gestations, or on what was documented at delivery. For the cohort study, gestational age at delivery was calculated based on the ultrasound examinations conducted by an experienced research sonographer, who used standardised assessment protocols and was blinded to the midwife GA assessment.

For all three study components, obstetric outcomes, including date of delivery and birthweight, were abstracted or digitised following delivery from the delivery record. Pregnancy loss was defined as any loss before delivery, with the composite variable including non-viable and ectopic pregnancy as determined by research sonographer; miscarriage was defined as loss <28 weeks of a clinically recognised pregnancy (27); stillbirths as fetal death occurring before or during labour/delivery (with 1-minute APGAR score of 0).

7.3.4 Covariates

For the PHDC and PER populations, data on risk factors for adverse birth outcomes were limited to data collected as part of routine care. For PHDC this included basic obstetric history and ART exposure history constructed from pharmacy dispensing data. While for PER this included more detailed data collected as part of routine care procedures such as body mass index (BMI) which was categorised as underweight (<18.5 kg/m^2), normal (18.5 - <25 kg/m^2), overweight (25 - 30 kg/m^2), obese (>30 kg/m^2) using weight at enrolment into ANC. Blood pressure was also measured at baseline as a part of routine care and categorised as normal (<120/80 mm Hg), prehypertensive (systolic 120–139 and diastolic 80–89 mm Hg Hg), and hypertensive (systolic ≥140 or diastolic ≥90 mm Hg) (28, 29). For the cohort study additional detailed data on possible risk factors for adverse birth outcomes were collected. This included a composite poverty score, based on current employment, housing type and access to household assets. This score was used as a proxy for socioeconomic status (SES) and categorised women as 'lowest', 'medium' or 'highest'(30). Alcohol use was measured using the 3-item Alcohol Use Disorders Identification Test-Consumption (AUDIT-C; range 0-12). For women, an AUDIT-C score ≥3 indicates hazardous drinking in the previous 12 months (31). Depressive symptoms were measured using the Edinburgh Postnatal Depression Scale (EPDS)(32), which has been validated for use in pregnancy (33). Our research group has previously used EPDS as a measure of self-reported depressive symptoms, rather than as a diagnostic tool (34). Threshold values of ≥13 indicated elevated depressive symptoms (32).

7.3.5 Statistical Analyses

Statistical analyses were performed using STATA version 14.0 (Stata Corporation, College Station, TX, USA), with the three data sources analysed separately. PHDC analyses were restricted to "probable pregnancies" classified based on having at least one high confidence event considered to be indicative of pregnancy (Rhesus test, pregnancy test, antenatal visit, maternal discharge summary), or combined moderate-confidence events (registration on mobile health messaging service (MomConnect)(35), antenatal booking visit, diagnosis or procedure codes indicating pregnancy, pap

smear test with pregnancy duration). These probably pregnancies had a pregnancy confidence score of 0.75 or greater (24).

Analyses across the three data source populations focused on two comparisons: women living with HIV vs those living without HIV, and among women living with HIV those initiating ART before pregnancy vs those initiating ART during pregnancy. The PHDC had an additional subset of untreated women living with HIV included in the ART status comparison. Maternal baseline characteristics and birth outcomes in each of the study components (PHDC, PER and cohort study) were stratified by HIV status and timing of ART initiation, with comparisons of proportions by Wilcoxon test, χ^2 test or Fisher exact test as appropriate. Associations between HIV/ART status with PTD were examined using modified Poisson regression with robust error variance. Because of the different levels and quality of data, the adjustment variables for each of the cohorts varied. Consequently, to facilitate a direct comparison, an additional analysis was conducted with adjustment by variables that were present in all three data sources.

7.3.6 Ethical Approvals

The study was reviewed and approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee and the Western Cape Department of Health. The PER and PHDC are resources of the Western Cape Provincial Government, as such they fall within its ethical and legal authority. The relevant PHDC and PER datasets were requested and issued to the investigators under the study number with no identifiers included. Written informed consent was obtained for study participation in the cohort study and included permission to abstract data from routine clinical records as well as permission to access linked electronic health data through the pregnancy and post-partum period. As data were collected through routine health services for the PHDC and PER, anonymised data was obtained so informed consent was not sought.

7.4 Results

7.4.1 Baseline characteristics

During the study enrolment period, 183 593 pregnant women aged ≥18 years formed the PHDC population following exclusion of duplicate, moderate and low confidence pregnancy events. Of these 148 593 were living without HIV (81%) and 35 353 living with HIV (19%) (Figure 7.2a).

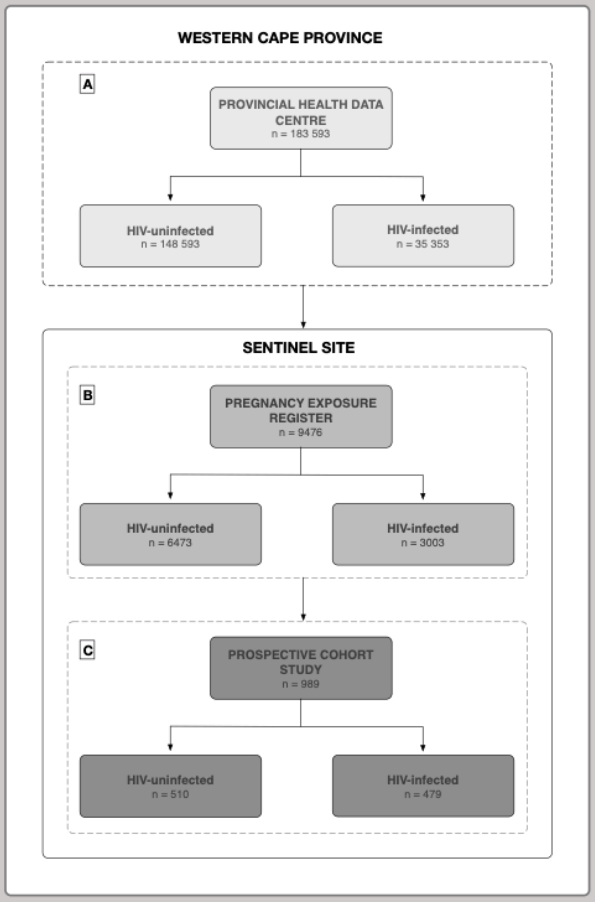

Figure 7.2 Relationship between three data sources
(A) provincial health data centre, **(B)** pregnancy exposure registry and, **(C)** cohort study

Figure 7.2 Relationship between three data sources

Among those living with HIV, 1859 (5%) were classified as untreated as they did not have electronic evidence of an ART start date at any point before delivery. Women living with HIV were older and less likely to be primigravid than women living without HIV, which was driven primarily by those who initiated ART before pregnancy (Table 7.1).

Table 7.1 Baseline characteristics of pregnant women (>18 years) in the provincial health data centre population (n=183 593)

	Total N = 183 593	HIV-uninfected N = 148 593	HIV-infected N= 35 353	HIV-infected N= 35 353		
				No ART initiated N=1 634	ART initiated before pregnancy N=19 842	ART initiated during pregnancy N=13 877
Age, years						
<24	67 410 (37)	59549 (40)	7861 (22)	465 (28)	3035 (15)	4361 (31)
25-29	52 646 (29)	42 119 (28)	10527 (30)	494 (30)	5777 (29)	4256 (31)
>30	63 890 (35)	46 925 (32)	16965 (48)	675 (41)	11 030 (56)	5260 (38)
Median, IQR	27 (23-31)	26 (22-31)	29 (25-33)	28 (24-33)	30 (26-34)	27 (23-32)
Gravidity						
1	97 925 (53)	81 144 (55)	16 781 (48)	955 (58)	7343 (37)	8483 (61)
2	56 376 (31)	44 644 (30)	11 732 (33)	471 (29)	7533 (38)	3728 (27)
≥3	29 644 (16)	22 804 (15)	6840 (19)	208 (13)	4966 (25)	1666 (12)
Missing	*1 (0)*	*1 (0)*	*0*	*0*	*0*	*0*
Median, IQR	*1 (1-2)*	*1 (1-2)*	*2 (1-2)*	*1 (1-2)*	*2 (1-2)*	*1 (1-2)*
Parity						
0	109 335 (59)	90 086 (61)	19 249 (54)	1083 (66)	9210 (46)	8956 (65)
1	49 051 (27)	38 767 (26)	10 284 (29)	374 (23)	6458 (33)	3452 (25)
≥2	25 560 (14)	19 740 (13)	5820 16)	177 (11)	4174 (21)	1469 (11)
Median, IQR	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)
Tuberculosis*						
Yes	1131 (0.61)	563 (0.4)	568 (2)	36 (2)	336 (2)	196 (1)
Isoniazid preventive therapy*						
Yes	6512 (4)	62 (0.04)	6450 (18)	58 (4)	3608 (18)	2784 (20)
Diabetes*						
Yes	6432 (4)	5527 (4)	905 (3)	34 (2)	587 (3)	284 (2)
Drug exposure (teratogen)						
Yes	17 057 (9)	11 826 (8)	5231 (15)	153 (9)	3415 (17)	1663 (12)
CD4, cells/mm³						
<200	3258 (9)	-	3258 (9)	180 (11)	1441 (7)	1637 (12)
201-350	5412 (15)	-	5412 (15)	274 (17)	2208 (11)	2930 (21)
351-500	5764 (16)	-	5764 (16)	221 (14)	2614 (13)	2929 (21)
>500	8469 (24)	-	8469 (24)	333 (20)	4609 (23)	3527 (25)
Median, IQR	422 (273-591)		422 (273-591)	387 (246-575)	455 (290-624)	397 (262-556)
Missing	*12450 (35)*		*12 450 (35)*	*626 (38)*	*8970 (45)*	*2854 (21)*
Viral load, copies/ml						
<400	3702 (10)		3702 (11)	134 (8)	2016 (10)	1552 (11)
401-1000	1502 (4)		1502 (4)	51 (3)	820 (4)	631 (4)
>1000	4569 (13)		4569 (13)	225 (14)	2236 (11)	2108 (15)
Median, IQR	806 (196-13980)		805 (196-13980)	1614 (256-22391)	693 (187-12331)	937 (201-15291)
Missing	*25580 (72)*		*25580 (72)*	*1224 (75)*	*14 770 (74)*	*9586 (69)*

* In pregnancy
ART – antiretroviral therapy

The PER population included 9476 women and of these 6473 were living without HIV (68%) and 3003 (32%) living with HIV (Figure 7.1b). Maternal age and gravidity were similar across the exposure groups as seen in the PHDC (Table 7.2).

Table 7.2 Baseline characteristics of pregnant women (>18years) in the pregnancy exposure registry population (n=9476)

	Total N=9476	**HIV-uninfected N=6473**	**HIV-infected N=3003**	**HIV-infected N= 3003**	
				ART initiated before pregnancy N=1986	**ART initiated during pregnancy N=1017**
Age, years					
<24	3334 (35)	2690 (42)	644 (21)	352 (18)	292 (29)
25-29	2741 (29)	1848 (29)	893 (30)	545 (27)	348 (34)
>30	3401 (36)	1935 (30)	1466 (49)	1089 (55)	377 (37)
Median, IQR	27 (23-32)	26 (22-31)	29 (25-34)	30 (26-34)	28 (24-32)
Gravidity					
1	2401 (25)	1904 (29)	497 (17)	261 (13)	236 (23)
2	3039 (32)	2051 (32)	988 (33)	617 (31)	371 (36)
≥3	3973 (42)	2486 (38)	1487 (50)	1094 (55)	393 (39)
Missing	*63 (1)*	*32 (1)*	*31 (1)*	*14 (1)*	*17 (2)*
Median, IQR	*2 (1-3)*	*2 (1-3)*	*2 (2-3)*	*3 (2-3)*	*2 (2-3)*
Parity					
0	1914 (20)	1383 (21)	531 (18)	307 (15)	224 (22)
1	3256 (34)	2172 (34)	1084 (36)	682 (34)	402 (40)
≥2	4188 (44)	2842 (41)	1346 (45)	974 (49)	372 (37)
Missing	*118 (1)*	*76 (1)*	*43 (1)*	*23 (1)*	*19 (2)*
Median, IQR	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)
Height, cm					
≤155	2538 (27)	1800 (28)	738 (25)	450 (23)	288 (28)
156-161	2589 (27)	1807 (28)	782 (26)	497 (25)	285 (28)
≥162	1886 (20)	1366 (21)	520 (17)	339 (17)	181 (18)
Missing	*2463 (26)*	*1500 (23)*	*963 (32)*	*700 (35)*	*263 (26)*
Median, IQR	158 (153-162)	158 (153-162)	158 (153-162)	158 (154-162)	157 (153-161)
Body mass index, kg/m^2*					
Underweight (≤18.5)	132 (1)	123 (2)	9 (0.30)	5 (0.2)	4 (0.4)
Normal (18.6-24.9)	1640 (17)	1231 (19)	409 (14)	249 (13)	160 (16)
Overweight (25.0-29.9)	1781 (19)	1205 (19)	576 (19)	350 (18)	226 (22)
Obese (≥30.0)	3172 (33)	2202 (34)	970 (32)	636 (32)	334 (33)
Missing	*2751 (29)*	*1712 (26)*	*1039 (35)*	*746 (37)*	*293 (29)*
Median, IQR	29 (24-35)	29 (24-35)	30 (26-35)	30 (26-35)	29 (25-35)
Wasting (<22 cm MUAC)	412 (4)	339 (5)	73 (2)	42 (2)	31 (3)
Median, IQR	87 (57-122)	87 (51-122)	96 (68-122)	96 (68-125)	87 (57-122)
Blood pressure, mmHG*					
Normal (Sys <120 or Dia <80)	4908 (52)	3192 (49)	1716 (57)	1167 (59)	549 (54)
Prehypertensive (Sys 120-139 and Dia 80-89)	3416 (36)	2448 (38)	968 (32)	630 (32)	338 (33)
Hypertensive (Sys >140 or Dia >90)	741 (8)	570 (8)	171 (6)	97 (5)	74 (7)
Missing	*411 (4)*	*263 (4)*	*148 (5)*	*92 (5)*	*56 (6)*

ART – antiretroviral therapy, MUAC – mid upper arm circumference, Sys – systolic, Dia - diastolic

*1st antenatal care visit

The prospective cohort study enrolled 989 pregnant women, with similar proportions of living without HIV (n=510, 52%) and living with HIV (n=479, 48%) women (Figure 7.1c, Table 7.3). Compared to women living without HIV, women who were living with HIV were older, less likely to be primigravid, less educated and of a lower socioeconomic status, and more likely to have mild anaemia. Gestational age at first ANC visit, BMI, blood pressure at booking, employment status, depressive symptoms and hazardous drinking did not vary systematically by HIV status (Table 7.3).

Table 7.3 Baseline characteristics of pregnant women (>18years) enrolled in the cohort study (n=989)

	Total N=989	**HIV-uninfected N=510**	**HIV-infected N=479**	**HIV-infected**	
				ART initiated before pregnancy N=337	**ART initiated during pregnancy N=142**
Age, years					
<24	267 (27)	185 (36)	82 (17)	44 (13)	38 (27)
25-29	294 (30)	161 (32)	133 (28)	82 (24)	51 (36)
>30	428 (43)	164 (32)	264 (55)	211 (63)	53 (37)
Median, IQR	28 (24-33)	26 (23-32)	30 (26-35)	32 (27-35)	27 (24-32)
Education (finished high school)	343 (35)	204 (40)	139 (29)	93 (28)	46 (32)
Employment status					
Employed	423 (43)	221 (43)	202 (42)	135 (40)	67 (47)
Socioeconomic status					
Lowest	325 (33)	146 (29)	179 (37)	135 (40)	44 (31)
Medium	281 (28)	147 (29)	134 (28)	84 (25)	50 (35)
Highest	381 (39)	215 (42)	166 (35)	118 (35)	48 (34)
Missing	2 (0.2)	2 (0.4)	0	0	0
Gravidity					
1	205 (21)	136 (27)	69 (14)	34 (10)	35 (25)
2	287 (29)	161 (32)	126 (26)	76 (23)	50 (35)
≥3	497 (50)	213 (42)	284 (59)	227 (67)	57 (40)
Median, IQR	*3 (2-3)*	*2 (1-3)*	*3 (2-3)*	*3 (2-4)*	*2 (2-3)*
Parity					
0	251 (25)	168 (33)	83 (17)	43 (13)	40 (28)
1	331 (33)	169 (33)	162 (34)	101 (30)	61 (43)
≥2	407 (41)	173 (34)	234 (49)	193 (57)	41 (29)
Median, IQR	1 (0-2)	2 (1-3)	1 (1-2)	3 (2-4)	2 (2-3)
Body mass index, kg/m^2*					
Normal (≤24.9)	163 (16)	84 (16)	79 (16)	54 (16)	25 (18)
Overweight (25.0-29.9)	260 (26)	127 (25)	133 (28)	98 (29)	35 (25)
Obese (≥30.0)	473 (45)	246 (48)	227 (47)	160 (47)	67 (47)
Missing	93 (9)	53 (10)	40 (8)	25 (7)	15 (11)
Median, IQR	30 (26-36)	31 (26-37)	30 (26-35)	30 (26-35)	30 (26-35)
Wasting (MUAC <22 cm)	18 (2)	10 (2)	8 (2)	5 (2)	3 (2)
Median, IQR					
Gestational age, median, IQR*ε	20 (14-25)	21 (15-26)	19 (13-23)	20 (14-24)	17 (12-22)
Singleton pregnancy	971 (98)	504 (99)	467 (97)	329 (98)	138 (97)
Blood pressure, mmHG*					
Normal (Sys <120 or Dia <80)	617 (62)	305 (60)	312 (65)	227 (67)	85 (60)
Prehypertensive (Sys 120-139 and Dia 80-89)	271 (27)	153 (30)	118 (25)	80 (24)	38 (27)
Hypertensive (Sys >140 or Dia >90)	39 (4)	17 (3)	22 (5)	13 (4)	9 (6)
Missing	*62 (6)*	*35 (7)*	*27 (6)*	*17 (5)*	*10 (7)*

ART – antiretroviral therapy, MUAC – mid upper arm circumference, Sys – systolic, Dia – diastolic
*completed weeks
* measured 1st antenatal care visit

Table 7.3 Baseline characteristics of pregnant women (>18years) enrolled in the cohort study (n=989) cont.

	Total N=989	HIV-uninfected N=510	HIV-infected N=479	HIV-infected	
				ART initiated before pregnancy N=337	ART initiated before pregnancy N=337
Haemoglobin, g/dl*					
Normal (≥11.0)	399 (40)	216 (42)	183 (38)	128 (38)	55 (39)
Mild anaemia (9.0-10.9)	233 (24)	105 (21)	128 (27)	94 (28)	34 (24)
Moderate/severe anaemia (<8.9)	37 (4)	20 (4)	17 (4)	13 (4)	4 (3)
Missing	*320 (32)*	*169 (33)*	*151 (32)*	*102 (30)*	*49 (35)*
Previous preterm delivery	19 (2)	11 (2)	8 (2)	5 (2)	3 (2)
Alcohol use§					
Hazardous drinking	*34 (7)*	*37 (7)*	*34 (7)*	*21 (6)*	*13 (9)*
Perinatal depression‡					
Depressive symptoms	*129 (13)*	*66 (13)*	*63 (14)*	*42 (13)*	*21 (16)*
Current ART regimen					
TDF-3TC-EFV	417 (87)	-	417 (87)	277 (82)	140 (98)
TDF-3TC-NVP	14 (3)	-	14 (3)	13 (4)	1 (1)
Other NNRTI-based	4 (1)	-	4 (1)	4 (1)	0
PI-based	24 (5)	-	24 (5)	23 (7)	1 (1)
Missing	*20 (4)*	-	*20 (4)*	*20 (6)*	*0*

TDF – tenofovir, 3TC – lamivudine, EFV – efavirenz, NVP- nevirapine, PI – protease inhibitor, NNRTI - Non-nucleoside reverse transcriptase inhibitors,
* measured 1st antenatal care visit
‡Based on Edinburgh Postnatal Depression Scale (range 0-30); a score of ≥13 indicates depressive symptoms.
§Based on the AUDIT-C (range 0-12); a score of ≥3 indicates hazardous drinking.

7.4.2 Birth outcomes

In the PHDC population, following exclusions of missing outcome data (n=7874, 4%) and multiple pregnancies (n=3728, 2%), 13% of pregnancies ended up in a loss (n=23 027), with the majority of losses due to pregnancy terminations (55%), other losses included miscarriages (34%) and stillbirths (11%) with no differences by HIV status. However, when assessed by ART status, women who initiated ART during pregnancy had a lower proportion of pregnancy terminations, stillbirths and miscarriages than women who either initiated before pregnancy or were untreated (Table 7.4). Among the 149 317 live singleton births, 14% were preterm with slight difference by HIV status (Table 7.4). The group of women who were untreated had the highest proportion of preterm deliveries (20%), however the majority of these were late and moderate preterm deliveries. It is possible that they were unable to initiate ART because they delivered preterm.

Table 7.4 Birth outcomes by HIV/ART status among pregnant women with singleton pregnancies in the provincial health data centre population (n=172 344)

Outcomes	**Total N=172 344**	**HIV-uninfected N=32 677**	**HIV-infected N=139 667**	**HIV-infected**		
				No ART initiated N=1 519	**ART initiated before pregnancy N=18 424**	**ART initiated during pregnancy N=12 734**
Pregnancy outcome						
Pregnancy loss	23 027 (13)	18 778 (13)	4249 (13)	353 (23)	2 819 (15)	1 077 (8)
Termination	12 683 (7)	10 453 (7)	2230 (7)	190 (12)	1538 (8)	502 (4)
Miscarriage	7755 (5)	6284 (5)	1471 (4)	133 (9)	993 (5)	345 (2)
Stillbirth	2574 (1)	2030 (1)	544 (1)	30 (2)	287 (2)	227 (2)
Other	12 (0.01)	8 (0.01)	4 (0.01)	0	1 (0.01)	3 (0.02)
Live births/delivered	149 317 (87)	120 889 (87)	28 428 (87)	1 166 (77)	15 605 (85)	11 657 (92)
Gestational duration (weeks)						
Post-term (≥42)	179 (0.1)	123 (0.1)	56 (0.2)	0	31 (0.2)	25 (0.2)
Term (37-41)	128 836 (86)	104 878 (87)	23 958 (84)	931 (80)	13 076 (84)	9951 (85)
Any preterm (gestational age <37)	20 302 (14)	15 888 (13)	4414 (16)	235 (20)	2 498 (16)	1 681 (14)
Late/Moderate preterm (32-37)	19 577 (13)	15 446 (12)	4131 (15)	228 (19)	2328 (15)	1575 (13)
Very preterm (28-31)	305 (0.2)	189 (0.2)	116 (0.4)	2 (0.2)	61 (0.3)	53 (0.5)
Extremely preterm (<28)	420 (0.3)	253 (0.2)	167 (0.6)	5 (0.4)	109 0.7()	53 (0.5)

ART – antiretroviral therapy

Adjusted PTD risk was higher in women living with HIV (ARR 1.15, 95% CI 1.11 - 1.1.18) (Table 7.5). Compared to women living with HIV, PTD risk in untreated women living with HIV was higher (ARR 1.53, 95% CI 1.37 - 1.72) than among those who were on preconception ART (ARR 1.15, 95% CI 1.11 - 1.20) and initiated during pregnancy (ARR 1.10, 95% CI 1.05 - 1.15) (Table 7.5). Among those who initiated ART, those who initiated during pregnancy were at lower risk of PTD compared to living without HIV reference group (Suppl. Table 9.5.1).

Table 7.5 Adjusted associations between HIV/ART exposure groups and preterm delivery among women with live singleton deliveries in the provincial health data centre population (n=149 317)

Outcome Measure	**Comparison A[†] (Ref category: HIV-uninfected)**			**Comparison B[††] (Ref category: HIV-uninfected)**		
		ARR [95% CI]	**P value**		**ARR [95% CI]**	**P value**
Preterm delivery (<37 weeks)*	HIV-infected	1.18 (1.15–1.22) [X] 1.15 (1.11-1.18) [XX]	<0.0001 <0.0001	No ART initiated[§]	1.53 (1.37-1.72) 1.51 (1.35-1.69)	<0.0001 <0.0001
				ART initiated before pregnancy	1.22 (1.17-1.27) 1.15 (1.11-1.20)	<0.0001 <0.0001
				ART initiated during pregnancy	1.10 (1.05-1.15) 1.10 (1.05-1.15)	<0.0001 <0.0001

ART – antiretroviral therapy, ARR – adjusted risk ratios
[†]Comparison A (HIV-uninfected vs HIV-infected), [††]Comparison B (HIV-uninfected vs HIV-infected - ART initiated during pregnancy vs HIV-infected - ART initiated during pregnancy (PHDC includes additional subset of women who did not initiate ART before delivery[§]))
* adjusted for age, gravidity, TB, smoking
[X] unadjusted analysis, [XX] adjusted analysis

In the PER population, 479 pregnancies (5%) had no outcome data with women living with HIV having a slightly higher proportion of missing data (6% vs 4%). Following exclusions of these missing outcomes and multiple pregnancies (n=57), 3% of pregnancies (n=254) ended up in a loss with the majority being miscarriages (63%) followed by stillbirths (23%). No differences were observed by HIV status or by timing of ART initiation (Table 7.6).

Table 7.6 Birth outcomes by HIV/ART status among pregnant women with singleton pregnancies in the pregnancy exposure registry population (n=8940)

Outcomes	**Total**	**HIV-uninfected**	**HIV-infected**	**HIV-infected**		
				No ART initiated	**ART initiated before pregnancy**	**ART initiated during pregnancy**
Pregnancy outcome	**N=8940**	**N=6152**	**N=2788**		**N=1845**	**N=943**
Pregnancy loss	254 (3)	155 (3)	99 (4)		65 (4)	34 (4)
Miscarriage	160 (2)	97 (2)	63 (2)		38 (2)	25 (3)
Stillbirth	58 (0.6)	38 (0.6)	20 (1)		16 (1)	4 (0.5)
Other	36 (0.4)	20 (0.4)	16 (1)		11 (1)	5 (0.5)
Live Birth	8686 (97)	5997 (97)	2788 (96)		1780 (96)	909 (96)
Gestational duration (weeks)						
Post-term (≥42)	26 (0.3)	14 (0.2)	12 (0.5)		7 (0.4)	5 (1)
Term (37-41)	7139 (82)	4940 (82)	2199 (82)		1467 (82)	732 (81)
Any preterm (<37)	1258 (15)	845 (14)	404 (15)		260 (15)	144 (16)
Late/Moderate preterm (32-37)	1002 (12)	675 (11)	327 (12)		211 (12)	116 (13)
Very preterm (28-31)	176 (2)	120 (2)	56 (2)		35 (2)	21 (2)
Extremely preterm (<28)	80 (1)	59 (1)	21 (1)		14 (1)	7 (1)
Missing	*263 (3)*	*189 (3)*	*74 (3)*		*46 (3)*	*28 (3)*

ART – antiretroviral therapy
* By design, the pregnancy exposure registry did not have women who did not initiate ART before delivery
Other – ectopic and/or terminations

Among 8686 live singleton births, 15% of deliveries were preterm, with no differences by HIV status (ARR 1.01, 95% CI 0.89 - 1.15) or ART status (those initiating ART before pregnancy ARR 0.96, 95% CI 0.83 - 1.11) and during pregnancy ARR 1.12, 95% CI 0.93 - 1.34) (Table 7.7).

Table 7.7 Adjusted associations between HIV/ART exposure groups and preterm delivery among women with live singleton deliveries in the pregnancy exposure registry (n=8686)

Outcome Measure	**Comparison A[†] (Ref category: HIV-uninfected)**			**Comparison B[††] (Ref category: HIV-uninfected)**		
		ARR [95% CI]	**P value**		**ARR [95% CI]**	**P value**
Preterm delivery (<37 weeks)*	HIV-infected	1.05 (0.90–1.16) [X]	0.353	ART initiated before pregnancy	1.02 (0.90-1.16) [X]	0.746
		1.01 (0.89-1.15) [XX]	0.841		0.96 (0.83-1.11) [XX]	0.601
				ART initiated during pregnancy	1.02 (0.90-1.16)[X]	0.186
					1.12 (0.93-1.34) [XX]	0.233

ART – antiretroviral therapy, ARR – adjusted risk ratios
[†]Comparison A (HIV-uninfected vs HIV-infected), [††]Comparison B (HIV-uninfected vs HIV-infected - ART initiated during pregnancy vs HIV-infected - ART initiated during pregnancy
*adjusted for age, gravidity, wasting
X crude xx adjusted

In the cohort study, following exclusions of multiple pregnancies (n=18, 2%) and missing data (n=20, 2%), 5% of pregnancies ended up in a loss (n=48). Of these the majority of losses were classified as stillbirths (46%) and miscarriages (40%) (Table 7.8). Among the 903 live singleton pregnancies, 11% of pregnancies were preterm with no differences by HIV and ART status.

Table 7.8 Birth outcomes by HIV/ART status among pregnant women with singleton pregnancies in the cohort study (n=951)

Outcomes	**Total**	**HIV-uninfected**	**HIV-infected**	**HIV-infected**		
				No ART initiated	**ART initiated before pregnancy**	**ART initiated during pregnancy**
Pregnancy outcome	**N=951**	**N=496**	**N=455**		**N=323**	**N=132**
Pregnancy loss	48 (5)	28 (6)	20 (4)		11 (4)	9 (7)
Miscarriage	19 (2)	12 (2)	7 (1)		5 (2)	2 (2)
Stillbirth	22 (2)	13 (3)	9 (2)		3 (1)	6 (4)
Other[§]	7 (1)	3 (1)	9 (1)		3 (1)	1 (1)
Live Births	903 (95)	468 (94)	435 (96)		312 (96)	123 (93)
Gestational duration (weeks)						
Post-term (≥42)	6 (1)	0	6 (1)		6 (2)	0
Term (37-41)	788 (87)	415 (89)	373 (87)		264 (84)	109 (89)
Any preterm (<37)	101 (11)	49 (10)	52 (12)		39 (13)	13 (10)
Late/Moderate preterm (32-37)	85 (9)	42 (9)	43 (10)		32 (10)	11 (8)
Very preterm (28-31)	13 (1)	6 (1)	7 (2)		6 (2)	1 (1)
Extremely preterm (<28)	3 (0.3)	1 (0.2)	2 (0.5)		1 (0.3)	1 (1)
Missing	*8 (1)*	*4 (1)*	*8 (1)*		*3 (1)*	*1 (1)*

ART – antiretroviral therapy

* By design, cohort study did not have women who did not initiate ART before delivery

[§]Other – ectopic (n=3) and terminations (n=4)

The lack of difference observed persisted in adjusted analyses, with no differences in the risk of PTD in HIV-infected women (ARR 0.95, 95% CI 0.61 - 1.47), or among those initiating before (ARR 0.93, 95% CI 0.57 - 1.52) and during pregnancy (ARR 0.99, 95% CI 0.52 - 1.87) (Table 7.9).

Table 7.9 Adjusted associations between HIV/ART exposure groups and preterm delivery among women with live singleton deliveries in the cohort study (n=903)

Outcome Measure		Comparison A† (Ref category: HIV-uninfected)			Comparison B†† (Ref category: HIV-uninfected)	
		ARR [95% CI]	P value		ARR [95% CI]	P value
Preterm delivery (<37 weeks)***	HIV-infected	1.14 (0.77–1.69) X	0.503	ART initiated before pregnancy	1.20 (0.78-1.82)	0.406
		0.95 (0.61-1.47) XX	0.551		0.93 (0.57-1.52)	0.769
				ART initiated during pregnancy	1.01 (0.55-1.86)	0.977
					0.99 (0.52-1.87)	0.973

ART – antiretroviral therapy, ARR – adjusted risk ratios
†Comparison A (HIV-uninfected vs HIV-infected), ††Comparison B (HIV-uninfected vs HIV-infected - ART initiated during pregnancy vs HIV-infected - ART initiated during pregnancy)
*** adjusted for age, gravidity, BMI, hypertension, hazardous drinking, depressive symptoms, PI regimen
X crude xx adjusted

7.4.3 Data source comparisons

Because of the differences in baseline characteristics available for the three populations it was only possible to assess differences in the distributions of the baseline characteristics age, gravidity and parity. Unsurprisingly the proportions by age group were similar between the PHDC and PER populations by HIV status, however in the cohort study where women were recruited based on numerous factors the proportions differed with an over-representation of women living with HIV who were aged >30 years. In contrast, for gravidity and parity the proportions by HIV status were similar between the PER population and cohort study with an under-representation of women who were primigravid and nulliparous compared to women in the PHDC population (Suppl. Table 9.5.2). Age is a risk factor for PTD, so we assessed overall PTD rates across the three data sources within each stratum of age and found that women in the cohort study had substantially lower rates of PTD than women from the PER and PHDC. When we stratified PTD by HIV status within age stratum, we found that women living without HIV age >30 years in the PER and cohort had a higher or similar proportion of PTD than their counterparts living with HIV suggesting that cohort selection could have played a part in attenuated findings in these two cohorts.

There was also a difference in the distribution of the different exposure groups. There was a significantly higher proportion of HIV-uninfected than women living with HIV in the PHDC (82% vs 18%) and PER (70% vs 30%) populations, reflecting the true population level distribution, than in the cohort study. By design there was a deliberate over-representation of women living with HIV (48%) and a corresponding under-representation of women living without HIV (52%) in the cohort (Suppl. Table 9.5.2). However, despite these differences in exposure group representation, the effect estimates for the association between HIV status and ART status and PTD were essentially the same in the PHDC, PER and cohort study with the main difference being the precision of the estimates (Table 7.3). When the effect estimates were adjusted for by the common variables age and gravidity, the estimates were also similar across the three data sources (Suppl. Table 9.5.3).

7.5 Discussion

Using data from a population level pregnancy surveillance, a pregnancy exposure registry and an observational cohort study, we evaluated the impact of cohort selection on internal validity and generalisability by comparing the association between HIV/ART status and PTD. Using data from the PHDC we found a 15% increased risk of PTD in women living with HIV compared to their counterparts living without HIV. The association was less pronounced in the pregnancy exposure registry and in the cohort study, with no significant differences seen by HIV status and timing of ART initiation. High-quality data on the safety and efficacy of most antiretrovirals in pregnancy is limited, because of the study designs or data sources typically used to generate evidence. Research on medications such as ART and other exposures in pregnancy is critical for public health, and so understanding the potential biases that may be seen in data from different sources is vital.

Electronic health data, such as PHDC data, are desirable because they are readily available and have a wide coverage of large portions of the target or underlying population, and thus provide useful public health information, even though they may lack in the number of variables collected (36). It has also been suggested that studies with larger sample sizes are more reliable than smaller studies because of their ability to detect clinically relevant differences, which can also increase the internal and external validity of findings (37). With the increased interest in “big data” analyses that incorporate data from large population groups and/or different data sources these larger sample sized studies are being advocated for (38). There are concerns around larger sample sizes since analysis power is substantially increased which could lead to an exaggerated tendency to reject null hypotheses with clinically negligible differences (37). This could be evident in our analyses, where

we found that the effect estimates between the three different sized populations were similar, with the effect only reaching statistical significance with the PHDC population. This is likely due to the large sample size in the PHDC, with narrow 95% confidence intervals and decreased variance of the estimates compared to the PER and cohort study. A recent cohort analysis conducted in the same research facility over a different time period, found no statistically significantly increased risk of PTD by either HIV status or timing of ART initiation (15). Other recent studies in similar settings also found similar findings (39, 40); likely because all these studies had smaller sample sizes, relative to the PHDC population. Sample sizes of the studies looking at this association now plays a more important role, and could serve as a possible explanation for findings, because differences in birth outcomes by HIV status have decreased in recent times. This is related to improvements in birth outcomes due to the increased coverage of maternal ART in particular the TDF+FTC+EFV regimen (41). Alongside this, the introduction of safer and efficacious regimens leading to higher and earlier ART uptake, has in turn improved the immunologic status of women living with HIV. Consequently, while these factors may explain part of the difference between the older and more recent studies, the main explanation for the difference seen here between PHDC and PER/cohort likely relates to the sample size. So significantly larger sample sizes will be required to investigate this association in the future, and this is likely to only come from the use of larger population level data from databases or registries, as seen here. Caution will however need to be exercised because the larger sample sizes may only allow more precise biased estimates (38).

Another possible explanation for the differences in our findings by data source could be related to bias. In the cohort study, we observed a lower-than-expected PTD rate and an attenuated, non-significant association between HIV status and PTD. This suggests that women recruited into the cohort study when compared to those accessing ANC in the same facility (PER) or those accessing ANC across the whole province (PHDC) appeared to be at inherently lower PTD risk. We speculate that these findings could in part be related to selective inclusion of healthier women in the PER and the cohort study arising from initial participation and self-selection. In the cohort study, where women were recruited and had to agree to be enrolled in the study, this selective inclusion may be related to self-selection with women more psychologically invested in their pregnancies more likely to enrol in studies and have better birth outcomes (42). While the catchment area for the women enrolled in the cohort study is a low-income area, a large proportion of women in the cohort study were classified as being of a higher socioeconomic status within that setting. Lower socioeconomic status is a well-known risk factor of adverse birth outcome so this could serve as another indication that the women enrolled in the cohort were at a relatively decreased risk of adverse outcomes (43). In the PER where all women accessing ANC in the facility were included, the selective inclusion

would be related to different factors than in the cohort study. The GMOU, which served as the sentinel site for the pregnancy registry, is a primary care facility which provides antenatal and obstetric care for low-risk pregnancies. Women with history of pregnancy complications or those requiring specialist review and/or interventions are referred to secondary or tertiary level obstetric facilities either at the first ANC visit or during pregnancy. Consequently, the women in the PER could also be a decreased risk of adverse birth outcomes.

Another important selection factor is the exclusion of women who had not initiated ART before they delivered from the group of women living with HIV in both the PER and cohort study does not give a true reflection of the underlying population. This group of women unsurprisingly had the highest PTD risk, consistent with findings from pre-ART era studies where HIV-infection was associated with increased PTD risk (44, 45). Their exclusion could have contributed to the attenuated associations, which were likely biased towards the null in the study cohort and PER when compared to PHDC. It should be noted that despite the initial participation and self-selection issues, which could have led to marginally biased estimates of the association between HIV/ART status, we found little indication of selection bias. We cannot, however, rule out the possibility of all or some of our findings being due to simple stochastic variation. We acknowledge that adjusted risk ratios have been presented despite the possibility that adjustment for potential confounders may remove some selection bias if the covariates are also associated with selection (46). While in general we do not have a good grasp of the predictors of pregnancy outcomes, making it difficult to know if we have adjusted appropriately, this scenario is possible in this analysis. Given that we adjusted for obstetric history even though women with a history of poor obstetric outcomes, by virtue of being referred out of the GMOU, could be excluded from the PER and cohort study. By adjusting for covariates that could influence selection the result could be closing of the open collider path that would otherwise induce a spurious association between HIV/ART and PTD (47, 48).

Despite the advantages of studies with larger sample sizes, prospective epidemiological studies which relate exposures or potential risk factors to the risk of the outcome of interest are still an important way of determining or investigating the etiology of outcomes (49). This is because robust measurement procedures can be used to measure exposure, risk factors at baseline so gold-standard data, determined by a definitive assessment, can be available for the whole study population (50). This may include important covariates for adjustment. This is evident in the cohort study where more detailed, complete and reliable information is collected on substance use, depression and ART regimen type, all which have been shown to be independently associated with PTD. Consistent, with findings from studies that include the protease inhibitor regimen (51-54), we

found that women living with HIV on this regimen were at almost three-fold the risk of delivering preterm compared to other, currently more commonly, used regimens namely TDF+FTC+EFV. We also found that women in the cohort study, who reported depression and/or hazardous drinking in the past 12 months, were also at increased risk of delivering preterm. Additionally, there appeared to be effect modification of the relationship between HIV status and PTD, with a higher incidence of PTD among the women living with HIV (compared to living without HIV) among the women exhibiting depressive symptoms. The use of prospective study designs also enables more accurate measurement of outcome incidence or precise timing of outcome events through observations at regular intervals over time can be achieved (55). In our cohort study we were able to ascertain a higher proportion of outcomes than with the other data sources. With electronic health data where data are collected for the purposes of the clinical encounter rather than being driven by research needs, the amount of time under observation (person-time) is determined by other factors, which could impact calculation of prevalence, incidence, and risk ratios (56).

A limitation of this analysis is that we were unable to assess and compare immortal time bias across the three data sources because the subset of women living with HIV who did not initiate ART was only present in the PHDC population not the PER and cohort study. Future studies investigating this association with this subset of women should utilise the proportional hazards regression modelling approach, treating ART initiation as a time dependent covariate which will enable partitioning of risk time relative to the events. The unique contribution of this study is the comparison of effect estimates between women enrolled in the cohort study and women accessing antenatal care at a sentinel site and across the Western Cape province, during the same time period. Many studies addressing the potential impact of cohort selection in epidemiologic studies usually assess difference in incidence estimates of exposures and outcome between participants and non-participants, and not the differences in the estimates of exposure-outcome associations. Although differences in incidence may point to the direction of non-representativeness, such findings are generally not adequate to reliably determine whether selection bias has affected the effect estimates. Further, through use the detailed data collection in the cohort study we were able to ascertain other potential modifying risk factors for PTD in our setting.

7.6 Conclusion

Despite the overrepresentation of the women living with HIV in the PER (to a lesser extent) and in the cohort study, bias in risk estimates appeared to be minimal. These findings also demonstrate the importance of including both population level data and prospective studies in future investigations of the association between HIV/ART status and adverse birth outcomes. Given the limited scope of clinical trials and cohort studies to detect safety signals in pregnant women, robust surveillance systems making use of routinely collected healthcare data at the broader population level will be essential to monitor trends of adverse birth outcomes. Simultaneous cohorts, in selected facilities, that use more robust measurements will also be necessary to improve understanding of the driving forces of the findings at population level. Additionally, as we move into an era of using big data to study medications in pregnancy, it will be important to not only focus on sample size but also on minimising bias through thoughtful study design, meticulous data collection, and use of appropriate statistical analyses.

7.7 References

1. Abrams EJ, Mofenson LM, Pozniak A, et al. Enhanced and Timely Investigation of ARVs for Use in Pregnant Women. *J Acquir Immune Defic Syndr*. 2021;86(5):607-615.
2. Matok I, Azoulay L, Yin H, Suissa S. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol.* 2014;100(9):658-662.
3. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol*. 2018;47(1):226-235.
4. White SW, Eastwood PR, Straker LM, et al. The Raine study had no evidence of significant perinatal selection bias after two decades of follow up: a longitudinal pregnancy cohort study. *BMC Pregnancy Childbirth*. 2017;17(1):1-10.
5. Snowden JM, Bovbjerg ML, Dissanayake M, Basso O. The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection. *Curr Epidemiol Rep*. 2018;5(4):379–387.
6. Weisskopf MG, Sparrow D, Hu H, Power MC. Biased exposure–health effect estimates from selection in cohort studies: are environmental studies at particular risk? *Environ Health Perspect*. 2015;123(11):1113-1122.
7. Biele G, Gustavson K, Czajkowski NO, et al. Bias from self selection and loss to follow-up in prospective cohort studies. *Eur J Epidemiol*. 2019;34(10):927-938.
8. Schnitzer ME, Blais L. Methods for the assessment of selection bias in drug safety during pregnancy studies using electronic medical data. *Pharmacol res perspect*. 2018;6(5):e00426.
9. Artama M, Gissler M, Malm H, Ritvanen A, Drugs and Pregnancy Study Group. Nationwide register-based surveillance system on drugs and pregnancy in Finland 1996–2006. *Pharmacoepidemiol Drug Saf*. 2011;20(7):729-738.
10. Howard TB, Tassinari MS, Feibus KB, Mathis LL, (Eds). Monitoring for teratogenic signals: pregnancy registries and surveillance methods. American Journal of Medical Genetics Part C: Seminars in Medical Genetics; 2011: Wiley Online Library.
11. Andrade SE, Raebel MA, Morse AN, et al. Use of prescription medications with a potential for fetal harm among pregnant women. *Pharmacoepidemiol Drug Saf*. 2006;15(8):546-554.
12. Andrade SE, Bérard A, Nordeng HM, et al. Administrative claims data versus augmented pregnancy data for the study of pharmaceutical treatments in pregnancy. *Curr Epidemiol Rep*. 2017;4(2):106-116.
13. Charlton RA, Neville AJ, Jordan S, et al. Healthcare databases in Europe for studying medicine use and safety during pregnancy. *Pharmacoepidemiol Drug Saf*. 2014;23(6):586-594.

14. Kukull WA, Ganguli M. Generalizability: the trees, the forest, and the low-hanging fruit. *Neurology.* 2012;78(23):1886-1891.
15. Malaba TR, Mukonda E, Matjila M, et al. Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age. *Paediatr Perinat Epidemiol*. 2022;36(4):525-535.
16. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. *Int J Epidemiol*. 2018;47(5):1714-1722.
17. Lash TL, Fox MP, MacLehose RF, et al. Good practices for quantitative bias analysis. *Int J Epidemiol.* 2014;43(6):1969-1985.
18. Nilsen RM, Surén P, Gunnes N, Alsaker ER, Bresnahan M, Hirtz D, et al. Analysis of self-selection bias in a population-based cohort study of autism spectrum disorders. *Paediatr Perinat Epidemiol*. 2013;27(6):553-563.
19. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology*. 2006:17(4):413-418.
20. Nilsen RM, Vollset SE, Gjessing HK, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol*. 2009;23(6):597-608.
21. Pizzi C, De Stavola B, Merletti F, et al. Sample selection and validity of exposure–disease association estimates in cohort studies. *J Epidemiol Community Health*. 2011;65(5):407-411.
22. Boulle A, Heekes A, Tiffin N, et al. Data centre profile: the provincial health data centre of the Western Cape Province, South Africa. *Int J Popul Data Sci*. 2019;4(2):1-11.
23. Mehta U, Heekes A, Kalk E, Boulle A. Assessing the value of Western Cape Provincial Government health administrative data and electronic pharmacy records in ascertaining medicine use during pregnancy. *S Afr Med J*. 2018;108(5):439-443.
24. Johnson LF, Mutemaringa T, Heekes A, Boulle A. Effect of HIV Infection and Antiretroviral Treatment on Pregnancy Rates in the Western Cape Province of South Africa. *J Infect Dis.* 2020;221(12):1953-1962.
25. Kalk E, Metha U, Slogrove AL, et al. Pregnancy Exposure Registry/Birth Defects Surveillance Programme In The Western Cape, South Africa: A Model For Low- & Middle-Income Countries. *18th Annual meeting International Society of Pharmacovigilance*, Geneva, Switzerland, 2018.
26. WHO. Preterm Birth. Geneva, Switzerland: World Health Organization; 2017.
27. WHO. Country, regional and global estimates. Geneva, Switzerland: World Health Organization; 2006.

28. Chobanian AV, Bakris GL, Black HR, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560-2571.
29. Brown MA, Magee LA, Kenny LC, et al. Hypertensive disorders of pregnancy: ISSHP classification, diagnosis, and management recommendations for international practice. *Hypertension*. 2018;72(1):24-43.
30. Brittain K, Mellins CA, Phillips TK, et al. Social Support, Stigma and Antenatal Depression Among HIV-Infected Pregnant Women in South Africa. *AIDS Behav.* 2017;21(1):274-282.
31. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med*. 1998;158(16):1789-1795.
32. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*. 1987;150(6):782-786.
33. Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS). *J Reprod Infant Psychol*. 1990;8(2):99-107.
34. Brittain K, Mellins CA, Phillips TK, et al. Social support, stigma and antenatal depression among HIV-infected pregnant women in South Africa. *AIDS Behav*. 2017;21(1):274-282.
35. Barron P, Peter J, LeFevre AE, et al. Mobile health messaging service and helpdesk for South African mothers (MomConnect): history, successes and challenges. *BMJ Glob Health*. 2018;3(Suppl 2):e000559.
36. Schneeweiss S. Learning from big health care data. *N Engl J Med*. 2014;370(23):2161-2163.
37. Faber J, Fonseca LM. How sample size influences research outcomes. *Dental Press J Orthod.* 2014;19(4):27-29.
38. Kaplan RM, Chambers DA, Glasgow RE. Big data and large sample size: a cautionary note on the potential for bias. *Clin Transl Sci*. 2014;7(4):342-346.
39. González R, Rupérez M, Sevene E, et al. Effects of HIV infection on maternal and neonatal health in southern Mozambique: A prospective cohort study after a decade of antiretroviral drugs roll out. *PLoS One*. 2017;12(6):e0178134.
40. Goldman G, Budhram S. A retrospective cohort study comparing pregnancy outcomes and neonatal characteristics between HIV-infected and HIV-non-infected mothers. *S Afr Med J*. 2020;110(6):502-504.

41. Moodley T, Moodley D, Sebitloane M, Maharaj N, Sartorius B. Improved pregnancy outcomes with increasing antiretroviral coverage in South Africa. *BMC Pregnancy Childbirth*. 2016;16(1):35-45.
42. Kramer MS, Wilkins R, Goulet L, et al. Investigating socio-economic disparities in preterm birth: Evidence for selective study participation and selection bias. *Paediatr Perinat Epidemiol*. 2009;23(4):301-309.
43. Ngandu CB, Momberg D, Magan A, et al. The association between household socio-economic status, maternal socio-demographic characteristics and adverse birth and infant growth outcomes in sub-Saharan Africa: a systematic review. *J Dev Orig Health Dis*. 2020;11(4):317-334.
44. Rollins NC, Coovadia HM, Bland RM, et al. Pregnancy outcomes in HIV-infected and uninfected women in rural and urban South Africa. *J Acquir Immune Defic Syndr*. 2007;44(3):321-328.
45. Brocklehurst P, French R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *Br J Obstet Gynaecol*. 1998;105(8):836-848.
46. Bliddal M, Liew Z, Pottegård A, et al. Examining nonparticipation in the maternal follow-up within the Danish National Birth Cohort. *Am J Epidemiol*. 2018;187(7):1511-1519.
47. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-625.
48. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39(2):417-420.
49. Phillips AN, Smith GD. The design of prospective epidemiological studies: more subjects or better measurements? *J Clin Epidemiol*. 1993;46(10):1203-1211.
50. Hubbard RA, Lett E, Ho GY, Chubak J. Characterizing bias due to differential exposure ascertainment in electronic health record data. *Health Serv Outcomes Res Methodol*. 2021;21(3):309-323.
51. Koss CA, Natureeba P, Plenty A, et al. Risk factors for preterm birth among HIV-infected pregnant Ugandan women randomized to lopinavir/ritonavir-or efavirenz-based antiretroviral therapy. *J Acquir Immune Defic Syndr*. 2014;67(2):128-135.
52. Powis KM, Shapiro RL. Protease inhibitors and adverse birth outcomes: is progesterone the missing piece to the puzzle? *J Infect Dis*. 2014;211(1):4-7.
53. Powis KM, Kitch D, Ogwu A, et al. Increased risk of preterm delivery among HIV-infected women randomized to protease versus nucleoside reverse transcriptase inhibitor-based HAART during pregnancy. *J Infect Dis*. 2011;204(4):506-514.

54. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.
55. Vassy JL, Ho Y-L, Honerlaw J, et al. Yield and bias in defining a cohort study baseline from electronic health record data. *J Biomed Inform*. 2018;78:54-59.
56. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health*. 2016;37:61-81.

Chapter 8: Discussion and recommendations

8.1 Introduction

Introduction of the triple-drug antiretroviral therapy (ART) during pregnancy, as an intervention to promote the health of women living with HIV and their infants, represents one of the most important achievements of the public health response to the HIV epidemic (1). Following 2013 WHO recommendations that all pregnant women living with HIV receive ART for life (Option B+), there was a significant increase in the global proportion of pregnant women receiving ART (2). This widespread use of ART in pregnancy led to significant reductions in new paediatric HIV infections as well as improvements in the long-term health of women living with HIV (3). In 2015, WHO recommended initiation of lifelong ART as soon as possible after diagnosis (universal test and treat), for all people living with HIV including pregnant women (4), leading to a substantial increase in the proportion of pregnant women receiving ART at the time of conception (5). The data from studies in this thesis relate to the previously recommended first line regimen (EFV+FTC+TDF); which was recently changed in the 2018 WHO treatment guidelines, with replacement of EFV with dolutegravir (DTG) (6).

Maternal ART use has been linked to with increased risk of preterm delivery (PTD), small-for-gestational age (SGA) and low birthweight (LBW) infants. This increased risk is concerning because these outcomes are major contributors to neonatal and child mortality and morbidity. Consequently, the association between maternal ART use and these adverse outcomes has been a prominent area of research in the field, especially in sub-Saharan Africa, where the vast majority of pregnant women living with HIV reside (7), and where rates of PTD, SGA and LBW are high. Most of the evidence on a possible association between ART in pregnancy and adverse outcomes comes from observational studies, despite their known limitations in establishing causality on the impact of exposures such as maternal HIV/ART on outcomes. Indeed, results from observational studies investigating this association are conflicting (8-14), with inconsistencies, particularly across the different income settings, possibly attributable to the stark baseline health and health-care disparities, and partly mediated through other health factors or factors related to low socioeconomic status (15-17).

A randomised control trial, PROMISE, showed ART exposure to be associated with adverse birth outcomes (18), however generalisability was affected by the use of ART regimens that are not commonly used across sub-Saharan Africa, and/or the use of monotherapy regimens as comparison groups. The Tsepamo study was conducted following the rollout of lifelong ART for all pregnant

women living with HIV using EFV+FTC+TDF (19). This study highlighted the importance of regimen type, with birth outcomes differing by regimen group. Women on the EFV-based regimens had the lowest proportion of adverse outcomes compared to women on the nevirapine -based and lopinavir/ritonavir-based regimens even when adjusted for age and other factors. More recently, the IMPAACT 2010/VESTED trial conducted to compare the safety of DTG and tenofovir alafenamide fumarate (TAF) containing ART initiated in pregnancy, encouragingly affirmed the current WHO recommendations to use DTG, with the newer DTG+FTC+TAF regimen showing a lower risk of adverse birth outcomes compared with either DTG+FTC+TDF or EFV+FTC+TDF regimens (20). Collectively these studies underscore the importance of the ongoing evaluation of birth outcomes with new ART regimens or HIV treatment options that are likely to be used in pregnancy.

Much of the literature on this association comes from studies conducted in LMICs, particularly sub-Saharan Africa. While health disparities may exist within sub-Saharan Africa across regions, within countries, and between different segments of the population, with improvements in PMTCT programmes the treatment and care profile of women living with HIV regionally and within countries has improved drastically everywhere. Consequently, any inconsistencies reported on the association with ART exposure in studies conducted in sub-Saharan Africa are now more likely to be linked to methodological issues. So going forward researchers investigating this association will need to think more deeply and critically about study design and analytical approaches used to ensure the studies are methodological sound. With this in mind, the overarching aim of this PhD research was to more reliable quantify the relationship between maternal ART use and adverse birth outcomes, by addressing the role of methodological factors inherent in observational research in this association. This chapter provides a discussion of key findings of this PhD thesis as a collective body of work, overall limitations, and recommendations for future research.

8.2 Discussion of key findings

Key methodological and clinical findings are summarised, divided into those related to the methodological factors and approaches explored (Table 8.1) and clinical findings related to the association between maternal ART use and adverse birth outcomes (Table 8.3). The PhD specific research objectives are addressed as appropriate (Chapter 1, Table 1.1), first in the methodological findings section and then in the clinical findings section. Throughout, preterm delivery is defined as delivery before 37 completed weeks of gestation, and small-for-gestational age as birthweight below the 10th centile for gestational age, and low birthweight as birthweight below 2500 grammes. The association of interest refers to the association between maternal ART use and adverse birth outcomes.

Table 8.1 Summary of key findings related to methodological approaches by research objective

Chapter and Objective		Summary of key findings

3	To identify the important methodological (selection and measurement bias) issues in the epidemiological study of preterm birth and its sequelae in the context of maternal infections and their treatment in LMICs. (Review Paper).	• **Selection bias**
		  • Study in Malawi found no difference in PTD by HIV status or timing of ART initiation
		  • Selection concern: postpartum recruitment leading to exclusion of high-risk pregnancies and infants could have biased estimates towards the null
		  • Systematic review of PTD by timing of ART initiation found increased risk with preconception ART
		  • Selection concern: exclusion of women delivering before ART initiation lowered risk estimates of women initiating during pregnancy
		• **Measurement bias**
		  • Studies in SA and Zambia found that gestational age assessment method impacted on estimates of risk of PTD
		  • Differences by assessment could be due to random measurement error
		  • Depending on timing of antenatal care presentation, LMP estimates errors could also be systematic

4	To examine the impact of measurement error related to gestational age assessment methods on the association between maternal ART use and preterm delivery.	• **Method of gestational age assessment impacted estimated PTD incidence:**
		  • PTD incidence using LMP-GA (36%) was double the incidence when using SFH-GA (17%) and triple when using US-GA (11%)
		  • Similar incidences were observed by HIV status using LMP-GA or SFH-GA; however, using US-GA PTD incidence was higher in women living with HIV (vs living without HIV)
		• **Method of gestational age assessment impacted observed PTD risk**
		  • In both unadjusted and adjusted analyses, increased PTD risk in living with HIV (vs living without HIV)women using US-GA (AOR1.95, 95% CI 1.10-3.46), risk in women living with HIV less pronounced with LMP/SFH-GA
		  • PTD risk was similar by timing of ART initiation for all methods
		• **Maternal factors influenced risk of gestational age under and over-estimation by both LMP and SFH (relative to US)**
		  • Elevated BMI and older age were associated with decreased risk of under-estimation
		  • HIV status and obesity were associated with increased risk of over-estimation by SFH

5	To examine the association between maternal ART use, timing of ART initiation and adverse birth outcomes, with bias adjusted gestational age.	• **In bias analyses, observed association between HIV status and PTD differed by gestational age assessment method**
		  • Observed association similar and weak for SFH and ultrasound, but LMP showed increased PTD risk for women living with HIV
		• **Last menstrual period gestational age was identified to be biased**
		  • LMP measures corrected using multiple overimputation

Chapter and Objective		Summary of key findings

6	To describe the natural history of blood pressure throughout pregnancy and assess whether this mediates the association between ART use and adverse pregnancy outcomes.	• **Blood pressure assessed using longitudinal approach more informative than with cross-sectional approach**
		  • Identification of systolic and diastolic trajectory groups enabled identification of high risk women in the trajectory group classified as abnormal (likely containing women experiencing pre-eclampsia)
		• **Trajectory groups an appropriate analytical approach for assessing the association between HIV and blood pressure status in pregnancy**
		  • Approach identified that abnormal blood pressure trajectories could be associated with ART initiation

7	To determine whether different data sources in the same underlying population impact the associations between maternal ART use and preterm delivery.	• **Three data sources with different data availability**
		  • Provincial data based on all pregnant women in Western Cape (PHDC), pregnancy exposure registry (PER) in one antenatal clinic and a dedicated research cohort based in same clinic
		  • Province data covered all births but with limited maternal and pregnancy variables, register more data information with most information from cohort with smallest sample size
		• **Similar effect estimates by data source**
		  • Estimated risk ratio of PTD by HIV status similar with only PHDC reaching statistical significance with a large sample size
		  • Precision of estimates differed by data source, highest in PHDC followed by PER then cohort study
		• **Some differences in women enrolled could explain differences in estimated association between data sources**
		  • Women living with HIV in PHDC included women untreated before delivery
		  • Women of lower PTD risk could have been included in the pregnancy registry and cohort study populations
		  • Exclusion of high-risk women who were untreated throughout pregnancy
		  • Little indication of selection bias
		• **Data from both large population-based sources and smaller prospective cohort studies complementary**
		  • Larger sample size in provincial dataset useful for overall association and see trends, whereas more detailed information in registry and cohort allows association with other variables that may impact the overall association at the cost of sample size
		  • More detailed information collected on important covariates identified variables for adjustment that were independently associated with preterm delivery

8.2.1 Key findings related to methodological approach used

A key skill for epidemiologists and clinical decision-makers, when evaluating relationships between exposures and outcomes of public health significance, is the ability to weigh up the different sources of evidence. This is especially important because a false 'conflict' has been created between the proponents of randomised control trials (RCT) and observational studies, despite both study types often being complementary and able to provide evidence not supplied by the other. As such this section will focus on illustrating the value of observational data using findings from this thesis to challenge the idea of fixed hierarchies of design and the perception that RCTs are the panacea in the study of exposures such as antiretroviral therapy use in pregnancy. Linked to this, this section will also address the broad underlying themes which may impact investigations of the relationship between maternal ART use and adverse birth outcomes.

Firstly, the nuances of epidemiological investigations of pregnancy and its outcomes complicate the design and interpretation of studies, so decisions around how we decide to undertake research by way of appropriate choices in design and data sources, exposure and outcome ascertainment and analytical approaches are vital. Secondly, what we research is important so if outcomes, exposures and important covariates are not adequately measured, we cannot address their potential impacts on adverse birth outcomes. Chapters 4-7 reported on what has been measured most notably gestational age, hypertension, obesity and depressive symptoms. This thesis demonstrated that bias and other methodological factors could impact the validity of inferences and make findings from an otherwise persuasive study difficult to interpret at best and irrelevant at worst. This may preclude finding the true association between maternal ART use and adverse birth outcomes through an inaccurate estimate (under- or over-estimate). This thesis also highlighted that a full appreciation of the many challenges inherent in perinatal epidemiology is essential in order to understand key elements involved particularly since there is limited scope for the adjustment or correction of most forms of bias during the analysis phase.

8.2.1.1 Study Design

Determining and understanding the aetiology of, and factors associated with, adverse birth outcomes and any related biological mechanisms, requires evidence be generated from well-designed studies with high quality study procedures and exposure, outcome and covariate measurement. In clinical research settings a common paradigm exists with a strong emphasis on hierarchies of research design, so RCTs are seen as the best source of evidence. The common view is that these trials are the gold standard strategy for assessment and/or cause-effect research of therapeutic agents. Consequently, findings from a single RCT are considered to provide the "truth",

while findings from observational studies are typically viewed suspiciously. In the epidemiology field, while not always bereft of trials, evidence is predominantly generated from observational sources (21). Therefore the view is that while RCTs should be conducted whenever possible or appropriate, when this isn't the case, important evidence coming from other sources should not be dismissed. This is particularly important in pregnancy research because, as briefly discussed in Chapter 2, RCTs are not always possible with the association of interest because of the known challenges of conducting RCTs in pregnant women.

The scarcity of RCTs investigating the association of interest is fairly well established, however it was only recently quantified through a systematic review and network meta-analysis of clinical trials assessing maternal ART regimens and adverse birth outcomes (22). Despite all antiretroviral (ARV) drugs and drug combinations being eligible for inclusion, regardless of whether they were recommended in current treatment guidelines or not, only seven unique studies were included with 6285 pregnant women living with HIV. Despite the limitation of external validity of RCTs, these trials could form an important part of the evidence base informing policy.
Despite the limitation of external validity of RCTs, conducting these trials in pregnant women remains essential and will still form an important part of the evidence base informing policy. These trials provide a controlled environment in which the safety and efficacy of medications to be used in pregnant women or women with pregnancy intentions can be monitored. However, the main issue is that they are not conducted often enough, and so given the number of women living with HIV using ART in pregnancy, this low number of women included in RCTs raises questions regarding the ability of these trials alone to appropriately inform policy. This means that for the foreseeable future the majority of policy decisions will be based on evidence from observational studies, so there needs to be a shift in our way of thinking about the evidence produced by these studies. While the hierarchy of study designs is useful for providing a checklist for the evaluation of clinical studies, there also needs to be an understanding that rigid application of this hierarchy, when evaluating studies, is not always appropriate. The complexities of clinical research mean that the hierarchy could be overly simplistic, and efforts should be made to ensure that observational data is evaluated alongside trial data to inform policy.

Contemporary criticism of observational studies is likely based on features of weaker poorly designed observational studies mistakenly used to criticise all observational research. Conventional wisdom suggests observational studies consistently provide biased results in comparison to those of RCTs, without taking into account the observational study type or how well it has been conducted. This could be why discussions of the conflicting data in earlier studies of maternal ART use and

adverse birth outcomes are frequently ascribed to the use of data from observational studies with no further exploration of their methodological qualities. This is issue is not unique to perinatal epidemiology or to investigations of the association of interest, similar concerns have been raised in other fields. This has been most notable with nutritional epidemiology where the reliance on observational data for addressing etiologic questions has recently been criticised. This criticism has emerged because several questionably designed and poorly executed recent meta-analyses have reached erroneous or misleading conclusions, reigniting arguments over nutrition debates that have already been settled (23). It should however be noted that nutritional epidemiology, like perinatal epidemiology, has design and analysis issues unique to the field. A misunderstanding of these issues can result in non-constructive and at times naïve criticism of observational studies.

Contrary to prevailing beliefs, average results from well-designed and well-conducted observational studies do not systematically overestimate the magnitude of exposure-outcome associations reported in RCTs (24, 25). While this was not the focus of this thesis, we were however, by addressing some of the methodological issues of past observational studies investigating the association of interest, able to demonstrate in this thesis the value and importance of using different types of observational data generated through the robust measurement of exposures, outcomes and important covariates. In perinatal epidemiology and other epidemiology fields (for example nutritional and chronic disease epidemiology), many research questions related to aetiological hypotheses which cannot typically be answered by RCTs can be answered well through prospective research studies and/or through use of existing data collected from different sources as demonstrated in Chapter 7.

8.2.1.2 Data sources

Appropriate sample size is a vital component for reliable, reproducible, and valid results and so decisions around data sources need to take sample size requirements into consideration. A disadvantage of purposively sampled prospective cohort studies is that significant amounts of time and resources are typically required to get large sample sizes. Additionally, even if sufficient participants can be recruited into these prospective studies there are always concerns around attrition and losses-to-follow up. All these factors often lead to decreased statistical power, with possible selection bias impacting validity and study inferences (26). As the immunologic status of women living with HIV improves with earlier initiation of safer ART regimens, we hypothesize that differences in adverse birth outcomes by HIV and ART status will become more subtle. Consequently, future investigations will need to consider the correct sample sizes needed to rule out any relevant differences between the groups, relative to the benefit of new treatment options.

An advantage of the use of substantially larger sample sizes will be enhanced precision in estimating overall trends in outcomes (27). The impact of sample size was evident in Chapter 7, where the effect estimates between the different data sources were similar, with only the estimates based on a large sample size reaching statistical significance. Additionally, special attention will need to be paid to the clinical relevance of reaching statistical significance. For example, in investigations of new treatment options such as the use of long-acting drugs in pregnancy, we may be willing to accept a 2% increase in the risk of PTD, in which case we would not need to power for a 1% increase, in the attempt to see a statistical difference. It is also important to note that as researchers strive to achieve these larger sample sizes to address the association of interest – there is recognition that this feature is less valuable if the sample does not represent the population to which the results will be generalised or if key information is missing on a non-random basis.

Realistically, in order to achieve the kinds of sample sizes required, data sources other than purposively sampled prospective studies will need to be explored. It will, however, be necessary to ensure that larger sample sizes improve generalisability and not just precision otherwise there could be misleading scenarios when there are spurious estimates with tight confidence intervals (28). Generalisability and precision of estimates in population-based data sources were explored in the different data sources used in this thesis. One data source that will offer unique opportunities in studies investigating the association of interest is electronic health records (EHRs), as they will enable answering of research questions requiring large sample sizes or obtaining detailed data on hard-to-reach populations (29). Over an extensive period, linkage of the different types of records (e.g. routinely collected health care records and disease-specific registries) will ensure high statistical power, in so doing lessening methodological issues related to attrition, recall bias and losses-to-follow up (30, 31).

As discussed in Chapter 7, data from the Provincial Health Data Centre (PHDC) of the Western Cape Department of Health (32), generated evidence at population level of all pregnant women accessing healthcare services in the province. This data source provided a substantially larger sample size than other studies that typically investigate the association of interest. Additionally, this data source provided a high level of external validity through inclusion of important population subgroups, such as women of reproductive age or individuals with multimorbidity's, who are often under-represented in such studies (31). Evidence generated from this province wide data is essential for policy making because it provides guidance as to which population groups may find a particular intervention clinically meaningful. One such group identified In Chapter 7, was women who did not

initiate ART before they delivered. By design, these women were not included in the cohort study because enrolment into the study only took place among women who attended their first antenatal care visit at the Gugulethu midwife's obstetric unit (GMOU), where untreated women would have been immediately initiated on ART. In order to achieve UNAIDS 90-90-90 HIV treatment targets, identification of pregnant women living with HIV remaining untreated will remain essential in public health programmes (33, 34).

With the expansion of novel therapies and treatment options for widely prevalent diseases, such as HIV, the lack of robust safety data for many medications used in pregnancy is concerning (35). EHR databases from large national pregnancy exposure registries (PERs) (36, 37), in combination with administrative databases (38-40), are being increasingly used in observational research on pregnancy and pregnancy outcomes related to medication use. Data obtained from such a registry based at the GMOU (41), which enrolled all women attending antenatal care services at the facility, was used in Chapter 7. Despite similarities in overall estimates with the cohort study, the larger sample size in the PER unsurprisingly improved the precision of the estimates with narrower confidence intervals. However, this increased sample size was of relatively limited value when compared to that of the cohort study because generalisability was not improved, as both these data sources "enrolled" from the same underlying population that excluded high risk women who were transferred to secondary or tertiary levels of care. An advantage the PER had over the PHDC as a data source was that despite being parsimonious by design, its prospective nature enabled collection of additional exposure or risk factor data during pregnancy than is possible with the PHDC.

8.2.1.3 Exposure ascertainment and definitions

A key component of exposure ascertainment and definition in pregnancy, is the time window of exposure, because this window should reflect the period when the exposure is having its effects relevant to the outcome of interest (42). Consequently, a thoughtful approach is required for ascertainment of exposure timing in pregnancy because of the unique features of pregnancy that exist related to it being time-limited, frequently unplanned and initially undetected (43). Further complexity can be introduced to ascertainment because different study designs provide differing data quality because of compromises made often with respect to measurement quality. With observational data sources such as those included in this thesis, there is a reliance on routinely available data, which at times may include proxy indicators, to identify exposure status, unlike with clinical trials in which mechanisms exist to ensure good quality exposure status data as well as the collection of relevant characteristics of exposure (44). As highlighted in Chapter 2, exposure definition and the potential for exposure misclassification can be linked to research objectives.

When the exposure status is maternal HIV infection this is typically a binary classification of exposed versus unexposed. In the two cohort studies in this thesis, women were recruited from antenatal clinics where there is widespread and reliable antenatal HIV testing in our setting (45, 46), so any misclassification would be minimal and not likely to bias estimates.

When the exposure status is ART use, similar to other medication exposure, this is typically classified as a binary of exposed versus unexposed during the pregnancy or by trimester (43, 47). It should however be noted that this approach to classification is not always a reflection of real-world exposure patterns so exposure status misclassification is likely (43). These misclassification concerns are well established in perinatal pharmacoepidemiologic studies investigating medications in pregnancy such as antidepressants (48). Despite the various complexities associated with ART use, misclassification is rarely highlighted as a concern in studies. A considered approach is required for ART status definition and classification because as highlighted in Chapter 2, maternal ART use is a function of duration of use, regimen type used and adherence, with each independently contributing to the adverse birth outcome risk. Misclassification of any of these factors can result in incorrect exposure group assignment, which in turn could lead to incorrect inferences. While the first two components are subject to some degree of misclassification, they are at least important considerations for the majority of researchers investigating the association of interest. This is not the case with the third component adherence (a proxy measure of drug levels) which is critical in pregnancy, as missed ART doses may lead to virological failure, increasing the risk of adverse outcomes. In observational studies an assumption is often made in analysis that a women classified as being on ART is fully adherent, mostly likely linked to the difficulties of measuring adherence. In clinical trials adherence can be objectively monitored (e.g. pill counts or limited medication dispensing), while this is less straightforward with observational data where adherence is often measured through self-report measures or pharmacy dispensing records as was seen with the PER and PHDC data sources in Chapter 7. So in a non-adherent woman classified as ART exposed, based on these measures, misclassification would be a major concern if she experienced an adverse birth outcome which could be incorrectly attributed to ART exposure.

Another important aspect to consider during the development of exposure definitions, is whether there are any known or hypothesised biological mechanisms of action between the exposure and the outcome of interest (44). Growing evidence suggests that the risk of adverse birth outcomes varies by ART regimen (19), consequently exposure definitions could be tailored to specific ART regimens or ARV classes due to differences in the mechanisms of action of the various medications. The ability to define exposures based on these factors depends on the available data, which needs to

be taken into account creating an operational definition of exposure (44). Antenatal medication exposure is usually determined based on self-report, clinical records and/or electronic pharmacy data. Some data sources may be more informative for different types of medication by improving the ability to precisely characterise usage, while others may just be a reflection of prescriptions and/or dispensing records but not actual use. Concerns are frequently raised about the possibility of misclassification with self-report, however in our settings, self-report measures for ART regimen have been shown to be reliable (49). A factor associated with accurate recall of medication that likely contributed to this is that there tends to be better recall when there is recent and repeated use of a medication as is the case with daily ART. In Chapter 7, ART regimen data was not available for the PER and PHDC data at the time of analysis, while in the cohort study this data was available based on self-report. The availability of this data enabled a valuable comparison between protease inhibitor (PI)-based and efavirenz (EFV)-based regimens, which will be discussed further in the clinical findings section below.

8.2.1.4 Outcome definitions

The choice of outcomes of interest drives important study design considerations such as data sources, frequency and length of follow-up assessments in the study following initial exposure, and sample size (50). To ensure comparability between exposure groups with as little outcome misclassification (or measurement error) as possible, clear and objective definitions of the outcomes of interest are required. Perinatal outcomes, which can be influenced by *in utero* events and exposures, are primarily defined according to when they occur or are detected during pregnancy or perinatally (51).

Interpretation of outcome status may need to vary depending on the time period under consideration. Preterm delivery (Chapters 4-7), size for gestational age (Chapters 4-5) and low birth weight (Chapters 5-6), were defined according to standardised clinical definitions, so there are no concerns with unclear definitions which could lead to misclassification. Definitions for the pregnancy loss outcomes were less straightforward because of differences in limits of viability in different settings, as discussed in Chapter 2. This was especially the case with the distinction between miscarriage and stillbirth, where legal and administrative definitions differ, so for ease of comparison with other resource limited settings where HIV is most prevalent, a definition was used that would enable international comparability (28 weeks).

An important issue that relates to perinatal outcome definitions that is typically not considered by researchers is the differentiation between processes and outcomes, as exemplified by the frequent

use of caesarean section as an outcome. Caesarean section can either be a process by way of being a delivery method or it can be an outcome of intrapartum care (52). Within the context of our association of interest, caesarean section was used as an intermediate measure as it did not give an indication of health status. Rather, because PTD is a complex syndrome with varying phenotypes (53), in addition to the traditional classification by gestational age, in Chapter 6 we also used a phenotype-based classification (53, 54). Preterm delivery was classified as spontaneous or clinically indicated given considerable etiological heterogeneity.

While the definitions of the outcomes in this thesis did not vary, this was not the case with different measurement methods used which had differing degrees of accuracy, depending on the data source. Measurement of perinatal outcomes can be challenging, as the relevant denominators (and numerators) are not always available (55). For the estimation of the risk of outcomes that were of interest in this thesis, namely preterm delivery, pregnancy complications, and obstetric intervention, pregnancies can serve as an appropriate denominator with the important consideration of time (i.e. gestational age) (56). The importance of time is related in part to the windows of exposure vulnerability for many outcomes of interest, which in studies of medication in pregnancy, can be narrow and specific, prolonged or unknown (43). Additionally, as highlighted in Chapter 2, the ambiguous timing of events in pregnancy means that some outcomes are unobservable at the time of occurrence (51). This uncertainty around the timing of onset can complicate establishment of temporality of the exposure-outcome relation.

In perinatal epidemiology, one of the most important measures is gestational age but it is well established that it is difficult to assess accurately. Expectedly the main source of bias in studies of outcomes based on gestational age relates to the choice of gestational age assessment methods. So it is quite surprising that in studies investigating maternal ART use and adverse birth outcomes little to no attention is paid to the impact of gestational age measurement error on the association. Of particular concern is that in the high HIV prevalence areas, where the majority of studies investigating this association take place, the best high-quality measures and procedures are more often than not used. Consequently, it is critical that researchers understand the extent and sources of misclassification in outcome measurement in order to determine their potential impact on estimates of the association of interest. This is especially important because while it was previously only hypothesised, this thesis provided definitive evidence that the associations between maternal ART and preterm delivery and size-for-gestational age have in part been inconsistent because of the heterogeneity gestational age assessment methods employed. Findings in Chapter 4 demonstrated that decision-making processes around gestational assessment methods should involve

consideration of the measurement properties specific to the population and setting in which the measures are to be used (50), and that bias due to mismeasured gestational age should be considered as an alternative explanation for null findings.

8.2.1.5 Analytical approaches

To ensure appropriate analytical choices are made for observational studies that minimise the effects of confounding and bias, thoughtfulness during data collection and analysis is required. For the most part, the form of the primary outcome of interest (continuous variable, dichotomous or polychotomous categorical variable) is used to determine the analytical methods. However the unique features of pregnancy mean that conventional analyses are not always appropriate so other, more sophisticated methodologies, may be necessary (57). Failure to account for the methodological and analytical challenges of pregnancy could hinder interpretability of results and muddy actionable conclusions. Several analytical factors need to be considered when investigating the association of interest.

Comparison groups

Construction of different exposure groups can not only result in different conclusions to the same research question but can also result in the introduction of bias (58). The use of different and/or inappropriate exposure group comparisons in the past is another factor that has contributed to conflicting data in earlier studies on maternal ART use and adverse birth outcomes. It is therefore critical that definitions or groupings reflect actual exposure status with as little misclassification as possible. Unlike in the past, studies investigating the association of interest now typically use one or both of the following binary exposure groups: women living without HIV versus women living with HIV on ART (HIV status), or, among women living with HIV, ART initiated preconception versus ART initiated during pregnancy (ART status).

In this thesis, the two binary exposure groups of HIV status and ART status (timing of ART initiation) comparisons were used in Chapters 4 to 7 (Table 8.2). Misclassification of ART exposure status as either before or during pregnancy is possible because of a combination of factors which include immediate ART initiation under the universal test and treat HIV treatment strategy and the fact that a pregnancy can go undetected for weeks or months. Consequently ART initiated outside the context of antenatal care (ANC) services could be classified as being before pregnancy even though it could actually have been initiated during the early stages of an as yet undetected pregnancy. This misclassification could potentially underestimate any true associations between ART status and adverse birth outcomes. Additional exposure group classifications included in Chapters 5 and 7 were

the categorical comparison of women living without HIV versus women living with HIV initiating ART before pregnancy versus women living with HIV initiating ART during pregnancy. In Chapter 5, a further group included the expanded ART initiation in pregnancy by trimester (Table 8.2). The binary classification of ART status, including women untreated in pregnancy (compared to ART treated women), is typically not an appropriate comparison for etiological research on the impact of ART. Untreated HIV infected women are likely significantly immunocompromised, with an inherently increased underlying risk of adverse birth outcomes, affecting comparability. Inclusion of this group of untreated women, as part of the ART status comparison group in Chapter 7 is important for public health purposes where trends and patterns are important. Additionally, as highlighted in Chapter 2 inclusion of this group of women are needed to gain a greater understanding of immortal time bias in the context of this association of interest. Particularly since lack of consideration of this immortal time can confer spurious benefit to the treated group. Consequently, their exclusion from recent studies could also be contributing to the attenuated findings observed, as discussed in Chapter 7 and in the section above.

Table 8.2 HIV/ART exposure comparison groups by objective/chapter

Chapter	Living without HIV vs living with HIV	Living without HIV vs living with HIV (ART before) vs living with HIV (ART during)	Living with HIV (ART before) vs living with HIV (ART during)	Living without HIV vs living with HIV (ART before) vs living with HIV (ART during) by semester
4	✓		✓	
5	✓	✓	✓	✓
6	✓		✓	
7	✓	✓	✓	

Time varying exposures

As previously discussed, the binary exposure classification approach does not always reflect real-world exposure patterns because there is no distinction between exposure patterns that are constant or varying over time (58). With the pregnancy exposure approach (Figure 8.1A), there is an assumption of constant effect and a failure to account for the exposure variability throughout pregnancy (59). This approach could induce etiologically inadequate conclusions because of the introduction of bias if the exposure effect differs over time (60, 61). To account for the exposure effect varying over time (Figure 8.1B and 8.1C), a commonly used approach is the trimester-specific exposure averaging approach (61).

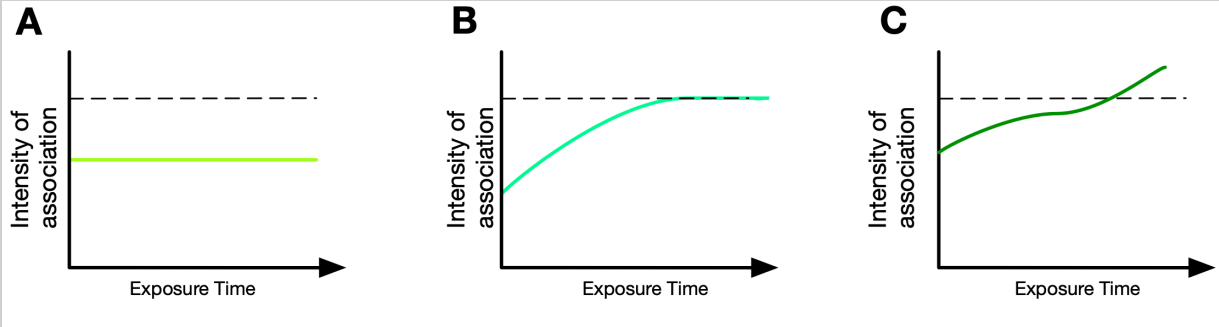

Figure 8.1 Examples of trajectories of association between an exposure to the subsequent outcome

(A) constant association, (B) time-varying with remote association or (C) time-varying with opposing remote and recent associations. The dashed horizontal lines represent no association

Adapted from: Wagner M, Grodstein F, Leffondre K, Samieri C, Proust-Lima C. Time-varying associations between an exposure history and a subsequent health outcome: a landmark approach to identify critical windows. *BMC Med Res Methodol.* 2021;21(266):1-15.

Alternative exposure modelling approaches exist that can summarise time-varying exposures that are not typically used in analyses of adverse birth outcomes (43, 58). Our primary exposure maternal ART use was not measured in a way that we could assess it's time varying nature i.e. different adherence patterns. So to explore an alternative approach to exposure modelling we used blood pressure (BP) as the main exposure in Chapter 6, which was modelled longitudinally in line with the natural history in pregnancy which typically decreases steadily until mid-pregnancy and then increases slightly again until delivery (62). Despite this pattern, BP is typically analysed using discrete, cross-sectional assessments at arbitrary time points (such as at the first ANC or enrolment visit) (63), with high BP defined by cut-off points used in non-pregnant adults. Using group-based trajectory modelling (GBTM), we were able to identify distinct trajectory groups which we classified according to the pattern of change throughout the antenatal period. When we assessed the association between BP and adverse birth outcomes, we found that the cross-sectional assessment approach, which classifies women as hypertensive using traditional cut-offs at first ANC, may not correctly identify women at increased adverse birth outcome risk. In contrast, GBTM enabled identification of women who were not classified as hypertensive but who were at increased risk, and thus the HDP and adverse birth outcomes association is likely to be more accurate using this approach. This analysis demonstrates that as researchers we need to think more critically about how we model exposure variables so that they can mirror real-world exposure patterns.

Bias Analyses

If the results of observational epidemiological studies are to be valid and applicable to decision-making about medication use in pregnancy, then it is of paramount importance that proper account be taken of biases to ensure that they are correctly controlled for (64). The current practices around

dealing with biases are at best acknowledgement and discussion of the possibility of bias impacting findings (typically in the limitations section of a manuscript) and at worst no mention at all. However no attempts are made to quantify biases through a series of sensitivity analyses to determine the possible direction and magnitude of any bias (65). In this thesis we applied a novel approach to ensure we had reliable gestational age estimates in the PIMS cohort where gestation was estimated using only LMP and SFH in some women. This was facilitated by the quantitative bias analysis we undertook to quantify the magnitude, direction, and effect of any bias induced by gestational age measurement error. Followed by multiple overimputation to address the bias in the LMP gestational age (Chapter 5). With observational studies involving both prospectively collected and medical record data, which both afford an excellent opportunity for research, more researchers investigating the association of interest in settings where ultrasound facilities are not readily available need to be encouraged to pay more attention to the potential impact of any bias that may be present.

8.2.1.7 Summary

The importance of studying perinatal health is matched by its unique challenges. Issues such as the extreme cohort attrition that occurs between conception and birth, the delimited nature of pregnancy duration, and accessing exposures and outcomes give rise to special cases of bias in perinatal epidemiology studies. The recommended principles discussed in this section will also be useful in understanding the strengths and limitations of new publications and for determining the risk of bias. Increasing the methodological rigour in the design and analysis of research on adverse birth outcomes will greatly improve the inferences that can be drawn from this work and ensure that future clinical and public health recommendations can be informed by the best possible estimates of the causal effects of maternal ART use on adverse birth outcomes.

In summary, the rigorous methodological approach applied to the observational data in this thesis, enabled reliable quantification of the association between maternal ART use and adverse birth outcomes findings of which will be presented in the next section. With a better understanding and quantification of errors introduced when using common error prone methods, the methodological findings in this thesis enabled identification and correction of potential biases in the association of interest.

8.2.2 Key clinical findings related to relationship between maternal ART use and adverse birth outcomes

8.2.2.1 Overview

With the programmatic successes of ART in pregnancy, leading to dramatic maternal health improvement and mother-to-child HIV transmission reduction, current concerns around ART exposure relate to the relationship with these adverse birth outcomes. This thesis set out to reliably examine and quantify the association between maternal ART use, timing of ART initiation and adverse birth outcomes, as well as any potential mediating effect of blood pressure on this association.

Over the past decade, study findings have suggested that HIV infection and ART use may be associated with an increased risk of adverse birth outcomes. However, a combination of methodological challenges, changing treatment guidelines and ART regimen, and sporadic clinical data has left the question unresolved, especially for pregnant women living with HIV in LMICs. This is also complicated by the fact that each new ARV drug in a new population requires a new investigation and hence a new question. In Chapter 3, the reflection of the important methodological considerations underlying perinatal epidemiological investigations that could apply to this association, highlighted that both the exposure and outcome could be significantly impacted by selection and measurement bias. Based on this, we addressed some of the methodological concerns in the methodological component of this thesis, in order to minimise and correct for any of the highlighted biases where possible. By ruling out the selection bias in the formation of the cohorts (Chapter 7) and accounting for outcome misclassification due to gestational age measurement error (Chapter 6), this thesis makes a substantial contribution by adding methodologically sound data to the literature around the association of interest in the context of WHO recommended treatment guidelines in the universal test and treat era. Because of the different aetiologies, the findings related to PTD, and fetal growth outcomes (SGA and LBW) will be discussed separately.

Table 8.3 Summary of key clinical findings related to relationship between maternal ART use and adverse birth outcomes by research objective

Chapter and Objective		Summary of key findings

5	Examine association between maternal ART use, timing of ART initiation and adverse birth outcomes, with bias adjusted gestational age.	• **Association between maternal HIV/ART status and preterm delivery**
		  • Risk of PTD did not differ by HIV status (ARR1.02, 95% CI 0.89-1.20)
		  • Risk of PTD by timing of ART initiation (compared to living without HIV), was similar in women initiating during pregnancy (ARR 1.07, 95% CI 0.89-1.31) or before pregnancy (ARR 0.97, 95% CI 0.79-1.18)
		  • Among women living with HIV, those initiating during pregnancy (compared to preconception initiation) did not differ in risk of preterm delivery (ARR 1.12, 95% CI 0.88-1.43)
		• **Association between maternal HIV/ART status and small for gestational age infants**
		  • Women living with HIV at increased risk of SGA (ARR1.37, 95% CI 1.10-1.72)
		  • Among women living with HIV, initiation in the second trimester associated with highest SGA risk (ARR1.53, 95% CI 1.11-2.12)
		• **Association between maternal HIV/ART status and low birth weight infants**
		  • Similar risks were observed by HIV status and timing of ART initiation

6	Describe natural history of blood pressure and assess whether this mediates association between ART use and adverse pregnancy outcomes.	• **Blood pressure trajectory groups associated with adverse birth outcomes**
		  • High normal (ARR 1.48, 95% CI 1.12-1.95) and abnormal groups (ARR3.18, 95% CI 2.32-4.37) associated with increased risk of PTD
		  • Abnormal groups (RR 2.81, 95% CI 1.90-4.15) associated with increased risk of LBW infants
		• **Effect modification**
		  • No effect modification by HIV status for PTD
		  • For LBW in the abnormal group, HIV status appeared to be an effect modifier with increased risk observed in uninfected women

7	Determine whether different data sources in the same underlying population impact the associations between maternal ART use and preterm delivery.	• **Association between maternal HIV/ART status and preterm delivery**
		  • Provincial health data centre
		  • Increased risk of PTD by HIV status (ARR 1.15, 95% CI 1.11-1.18), with highest risk among untreated HIV-infected women (ARR 1.53, 95% CI 1.37-1.72)
		  • Women living with HIV (compared to living without HIV), had increased PTD risk, which was similar whether initiation occurred preconception (ARR 1.15, 95% CI 1.11-1.20) or during pregnancy (ARR 1.10, 95% CI 1.05-1.15)
		  • Among those living with HIV, those initiating during pregnancy (compared to preconception initiation) were at decreased risk of preterm delivery (ARR 0.93, 95% CI 0.88-0.99), but this did not reach statistical significance
		  • Pregnancy exposure registry
		  • No difference observed in risk of PTD by HIV status (ARR 1.01, 95% CI 0.89-1.15)
		  • Risk of PTD by timing of ART initiation (compared to living without HIV), was higher in women initiating during pregnancy (ARR 1.12, 95% CI 0.93-1.34) than before pregnancy (ARR 0.96, 95% CI 0.83-1.11), but not statistically significant
		  • Women living with HIV initiating during pregnancy (compared to preconception initiation) were at increased risk of preterm delivery ARR 1.12, 95% CI 0.91-1.38)
		  • Cohort study
		  • No difference observed in risk of PTD by HIV status (ARR 0.95, 95% CI 0.61-1.47)
		  • Risk of PTD by timing of ART initiation (compared to living without HIV), slightly higher in women initiating during pregnancy (ARR 0.99, 95% CI 0.52-1.87) than before pregnancy (ARR 0.93, 95% CI 0.57-1.87)
		  • those initiating during pregnancy (compared to preconception initiation) were at non-significant increased risk of preterm delivery (ARR 1.08, 95% CI 0.53-2.18)

8.2.2.2 Association between maternal ART use and preterm delivery

Among the women enrolled into the two cohort studies (PIMS and B-Positive) and accessing antenatal care at the GMOU included in the registry (PER), we observed an overall PTD incidence which was not as high as we had anticipated. Alongside this, we found no appreciable difference in risk by HIV status and timing of ART initiation in the cohort studies, however in the provincial dataset both were associated with increased risk (Chapter 5 and 7, Table 8.3). There are numerous factors that likely contributed to the difference between the PTD findings presented in this thesis, and findings in the past. Previously, an increased risk in PTD was observed among women living with HIV in the pre-ART era (i.e. untreated) (66), as well as those on ART (vs women living without HIV), and among women initiating ART preconception (vs women initiating ART during pregnancy (67-71).

With the universal test and treat guidelines being implemented, ART is now initiated at HIV diagnosis regardless of pregnancy status. So women living with HIV accessing antenatal care services, during the course of the cohort studies included in this thesis, were all on lifelong ART initiated before or during pregnancy, with the safe and efficacious TDF+FTC+EFV regimen as the preferred regimen option. We hypothesise that both of these factors played a significant role in the improvements in maternal health and infant outcomes seen in recent times, including in this thesis. The finding of no statistically significant differences in risk of PTD by HIV status and by timing of ART initiation shown in this thesis, lends weight to speculation that important drivers of adverse birth outcomes in the past may not be as prevalent. These include HIV-related immunological characteristics, timing of treatment initiation and the use of ART regimens now less commonly used. Previous guidelines based treatment eligibility on CD4 cell count thresholds or WHO staging. So women eligible for treatment before they became pregnant because they were at advanced HIV stage. In more recent times, the pool of women initiating preconception has become diluted through the inclusion of women living with HIV who could be considered healthy and are on preconception ART because of current test and treat guidelines not because they were at advanced disease stage. This dilution means the group of women on preconception ART are likely to more closely resemble those initiating ART during pregnancy.

While there is no question that pregnant women living with HIV should receive ART, questions do remain around specific ART regimens and their association with increased risk of PTD. Before WHO recommended TDF+FTC+EFV as the first-line regimen, national guidelines varied across sub-Saharan Africa on the recommended regimens for use during pregnancy. At the time comparability of study findings were hampered because regimens used in some studies were not always widely used within the region. This was seen with the PROMISE study, which provided randomised trial data on the risks and benefits of triple drug ART, however the regimens compared were zidovudine-based ART

(zidovudine (AZT)+ lamivudine (3TC) +lopinavir–ritonavir (LPV/r)) and TDF-based ART (TDF+FTC+LPV/r) (18).

Conventional wisdom is that the use of a wide range of regimens used in studies such as PROMISE explains why findings on the impact of different ART regimens on adverse birth outcomes reported from different studies have been conflicting (22). However, a major issue linked to this school of thought is that ART is often been thought of as a single strategy, which is not appropriate, rather each ART regimen needs to be considered separately, and with this an understanding that results may vary by regimen comparison. An observational birth outcomes surveillance study in Botswana recently confirmed that specific ART regimens were associated with worse adverse birth outcomes (19), with PIs more commonly implicated in the occurrence of PTD than nucleoside and non-nucleoside reverse transcriptase inhibitors (NRTIs and NNRTIs). In Chapter 7, the impact of regimen type was demonstrated by the finding that the limited number of women living with HIV on a PI regimen were at almost three-fold the risk of delivering preterm compared to women using other, currently more commonly, used regimens namely TDF+FTC+EFV.

What is the significance of these findings?

The PTD findings in this thesis are an important contribution to understanding the current state of this association, because they demonstrate that concerns around HIV-infection and ART exposure in pregnancy remain valid as they still play a role in the occurrence of PTD, particularly when PI-based regimens are involved. Encouragingly the majority of preterm infants in our cohorts were classified as being appropriate-for-gestational age (AGA), and they are considered healthier than their SGA preterm counterparts and will likely experience catch-up growth that will put them on par with term infants within the first two years of life if they have adequate nutritional support (72). While this is encouraging for short and long term child health outcomes there is still an urgent need for additional research on among specific groups of women in our context who are at increased risk of delivering preterm and subsequent adverse outcomes. For instance, we found that women initiating ART later in pregnancy were at increased PTD risk which is of great concern because by virtue of them accessing ANC later in pregnancy they are already at increased risk of adverse maternal and infant outcomes. Public health approaches in collaboration with diverse stakeholders will be required to identify high risk women who need more attention during pregnancy which may involve transfers to higher levels of care.

Another important take home point from this research is that with future investigations of this association, a more circumspect approach to decisions around data sources and sample sizes will be necessary depending on the main objective. On one hand the objective could be to establish the

estimates of risk by HIV/ART status, or on the other establish potential modifiable risk factors. In the past, there was no question that a prospective cohort study was the most appropriate data source choice for both of these study objectives. However, when the goal is risk estimates by HIV/ART status this may no longer not always be the case because the differences in maternal health are now less stark than previously observed as mentioned in section 8.2.2.2. Consequently, for comparisons of PTD risk by HIV status and/or ART status, the smaller sized prospective studies that were typically used will now struggle to detect subtle but clinically relevant differences between these groups of women. As demonstrated in Chapter 7, detection of any potential increased PTD risk by HIV and/or ART status, will now require the use of significantly larger sample sized data sources such as the population-level EHRs.

However, when the goal is a more comprehensive exploration of risk factor, then prospective studies are still the best option because they enable detailed collection of risk factor data which could be useful in elucidating study findings. So for example with ART regimen types, specifically PIs, appearing to be associated with increased PTD risk, it will be essential to monitor other regimens using this study design as well as thoroughly investigating any potential drug mediated mechanisms as new treatment options are introduced through smaller mechanistic studies. Longitudinal data on blood pressure and maternal weight changes throughout pregnancy is another example of important risk factor data that could be collected.

Factors related to data sources, sample size and ART regimen will also need to be important considerations when evaluating findings from studies evaluating ART and other medication safety in pregnancy. Inaccurate conclusions from studies can have far-reaching consequences, as demonstrated by the fall-out from two important HIV treatment studies. First, based largely on the PROMISE trial findings, a panel of healthcare providers, researchers, and patient representatives issued controversial recommendations regarding ART in pregnancy, advocating for AZT+3TC-based regimens over a TDF+FTC containing regimen for pregnant women living with HIV (73), despite this regimen being less well tolerated. The guidelines committees and the PROMISE study team disagreed with these recommendations, with the PROMISE team, in particular, encouraging a biological reason for the increase in infant death to be determined before such assertions, this has never been found. Second, the early analysis of the Tsepamo study birth defects data has meant that despite a more holistic understanding of the risks versus benefits of DTG, following initial concerns, disparities in DTG uptake still exist affecting women of reproductive age. A recent study had demonstrated the extent and persistence of disparities despite WHO recommendations to expand DTG access (74). This could have implications for HIV treatment outcomes because of the

superior efficacy and tolerability of DTG (75) as well as its anticipated population health benefits (76).

So in summary, as new regimens and treatment options become available, an important contribution this thesis makes is the demonstration of the necessity of the inclusion of both prospective studies, which collect more detailed risk factor data, and large population level data sources, which can detect the smallest magnitude of the effect, for safety evaluations as they provide complementary data.

8.2.2.3 Association between maternal ART use and fetal growth outcomes

Among the women enrolled into the PIMS cohort study, we observed an increased risk of SGA by HIV status, but not by timing of ART initiation (Chapter 5, Table 8.3). While the LBW outcome is less informative, we evaluated it as a supplementary outcome for comparison purposes, because of the absence of reliable GA assessment in some studies conducted in LMICs precluding SGA from being diagnosed. We found no difference in the risk of LBW by HIV status or timing of ART initiation (Table 8.3). The highest proportion of LBW was observed among women initiating ART in the second trimester, which could give an indication of the aetiological pathway.

What is the significance of these findings?

The SGA findings in this thesis are a valuable addition to the body of evidence because studies exploring the occurrence of SGA within the context of maternal ART use are scarce. These findings also point out the need for more nuanced investigations of fetal growth outcomes. The limited data on SGA is largely due to the difficulties in studying fetal growth outcomes because they are a result of complex interactions between growth potential that is genetically pre-determined and numerous pregnancy-related factors (maternal, placental and fetal) (77). Furthermore, pinpointing infants at increased risk of perinatal morbidity or mortality because they are SGA is complicated because SGA infants can differ from growth-restricted infants principally because SGA also encompasses constitutionally small but healthy infants at lower risk of abnormal outcomes. As a result of these complexity, studies in LMICs have tended to focus on PTD as the outcome of interest. However as discussed in section 8.2.2.2, maternal health improvements and current ART regimens mean that as the incidence of PTD related to HIV infection and maternal ART use decreases, heightened focus will be required on SGA risk linked to maternal ART use. This is critical because SGA is clinically worse than PTD per se, as SGA infants are at significantly increased risk of child morbidity and mortality than AGA preterm infants (78-80).

Fetal growth disruptions leading to SGA and LBW infants are of particular concern because they are associated with both linear growth stunting and undernutrition among young children, which can have far reaching long-term health and developmental consequences (81). From our findings, it is unclear what the proportion of infants classified as SGA were growth restricted, and this is because like most clinical and epidemiological research investigating this outcome, we used birthweight-for-gestational-age references. The problem with this approach is these population references, which provide birthweight percentiles by each gestational week, tend to be developed using with very large mostly population-based databases (82). However, there robust evidence that shows that preterm infants are more likely to be growth-restricted (83-85), meaning that their birthweights are not representative of all infants at a given gestational week in preterm. Therefore, there is likely to be a significant under diagnosis of SGA infants among PTD infants when population birthweight references are used (82). Given that these infants are the ones of most concern, in future investigations of this association it will be essential to use methods that identify which infants have experienced some form of abnormal growth that is precluding them from meeting their growth potential and putting them at increased risk of morbidity or mortality.

In order to gather more robust epidemiological and mechanistic data on the association between maternal ART use and fetal growth outcomes such as SGA careful consideration will be required for how studies ascertain and classify outcomes. For starters investigations of SGA as an outcome should include data collected throughout pregnancy fetal growth patterns not just a definition based on size at delivery. Next classification system of fetal growth needs to be utilised which can accurately distinguish between normal and abnormal growth which is based on the risk of perinatal morbidity and mortality and/or life-course morbidity. Given that fetal size alone cannot accurately provide insight on whether an SGA infant is constitutionally small or growth-restricted, integration of other indicators of maternal, placental and fetal health will be also essential. Ideally a definition of growth restriction should consider fetal growth potential, current fetal size, maternal, placental and fetal health, and, if available, fetal growth velocity (82).

Lastly, while we did not find any increased LBW risk associated with the main exposures of interest, we did find an increased risk was associated with abnormal blood pressure trajectory over pregnancy. These findings reiterated the importance of ensuring that women with pre-eclampsia that is either new onset or superimposed on chronic hypertension have comprehensive plans of management with regular maternal review and strict blood pressure control and close surveillance of fetal growth and wellbeing, particularly in high HIV-settings.

8.2.2.4 What are the potential aetiological pathways for preterm delivery, small for gestational age and low birth weight infants linked to maternal ART use?

An understanding of the processes underlying adverse birth outcomes is essential because these may facilitate improved clinical management. However, the ongoing search to elucidate the underlying causes and pathophysiological mechanisms that exist as a result of HIV-infection and maternal ART use, is hampered by several challenges. First, the basic physiological and pathophysiological mechanisms that underlie the timing of onset of human parturition and adverse birth outcomes (PTD, SGA and LBW), respectively, are not yet well understood. Second, the study of infection processes in pregnancy is complicated by the effects that pregnancy-related alterations in maternal immunology.

Successful pregnancy, including optimal fetal growth, depends the immune tolerance and suppression processes being carefully balanced (Figure 8.2). The main mechanism operating concurrently to maintain pregnancy include shifts between pro-inflammatory and anti-inflammatory dominant environments during different phases (86-88), and the numerous placental processes which include progesterone production to inhibit uterine contractions (89).

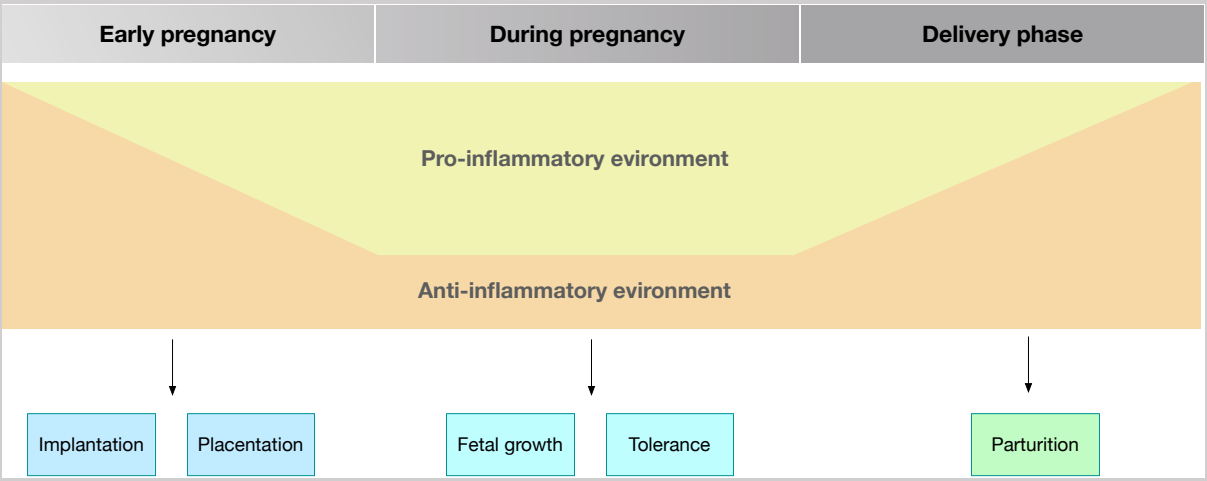

Figure 8.2 Immune environment during pregnancy

Adapted from: Förger F, Villiger PM. Immunological adaptations in pregnancy that modulate rheumatoid arthritis disease activity. *Nat Rev Rheumatol.* 2020;16(2):113-122.

Given the importance of the inflammatory profile, placental functioning and progesterone during pregnancy, we speculate that immune and placentation dysregulation could be linked to the aetiology of the adverse birth outcomes investigated in this thesis. The basis for our hypotheses is that HIV has been associated with chronic inflammation (90) and placental insufficiency features such as maternal vascular malperfusion (MVM) (91, 92). So irrespective of the mechanisms that

underlie inflammation, it is likely that associations observed between maternal ART use and adverse birth outcomes are driven to some extent by chronic inflammation. This is evidenced by inflammatory changes observed in the placentae of untreated and treated women living with HIV (93, 94).

Preterm delivery pathways

Across the analyses presented in this thesis, the highest proportions of PTD were found to be among untreated women and women initiating ART preconception (Chapter 7) as well as among women initiating in the 3rd trimester (Chapter 5). Put together these divergent findings suggest that in each of these groups of women different PTD aetiological pathways may be at play. Given that PTD, as discussed in Chapter 3, is a complex syndrome with multiple aetiologies that converge upon the final common pathway of early delivery (95), the findings of increased PTD proportions among these three groups of women are not surprising. The PTD classification is predominantly based on gestational age thresholds, however more nuanced phenotypic classifications exist which are based on physiological and pathophysiological pathways (53, 96). These other classification systems, which are focused on proposed aetiologies, are likely to be more useful for understanding and distinguishing between the pathways leading to PTD with initiation preconception and those leading to PTD with initiation in the third trimester.

We hypothesise that ART-mediated immune reconstitution was responsible for the majority of preterm deliveries we observed in the women initiating in pregnancy. This pathway has long been speculated to be linked to changes to T-helper (Th) cell environments when ART initiation occurs during pregnancy with an early shift back to a pro-inflammatory Th1 dominant environment (97). This disruption of the regular immune programme of pregnancy is not optimal for the maintenance of pregnancy, because it leads to increased levels of cytokines associated with early initiation of uterine contractions and thus an early delivery could occur (97-99). We also speculate that it is biologically plausible that among women initiating in the third trimester, the occurrence of the ART-mediated immune reconstitution at this time could be the reason for us observing the highest proportion of PTD in this trimester (Chapter 5). We suspect that this interference of the normal pregnancy environment could trigger early labour of an otherwise appropriately sized fetus that has is developmentally mature enough to survive post-delivery.

Given that ART results in viral replication control which in turn dramatically reduces inflammation and immune activation over time (100), we do not think that this pathway lead to increased risk of PTD among preconception ART initiators (Chapter 7). As highlighted in Chapter 3, selection bias has previously been put forward as a reason for increased risks associated with preconception ART

initiation (100, 101). Following investigation we excluded this as an explanation because the increased risk in preconception initiators was still observed (albeit at a reduced level) when the untreated women were grouped with those initiating during pregnancy. A recent analysis we conducted of a subset of women in the PIMS cohort with placentas collected at delivery, has helped cement our belief that the increased risk of PTD in preconception initiators is linked to ART exposure during placental implantation and development. We found an increased risk of MVM in placentas from women who had initiated ART preconception, but not those initiating during pregnancy (102). In turn, MVM was associated with increased risk of PTD (102).

Lastly, another pathway that could contribute to increased risks in preconception initiators is the use of ART regimens associated with increased PTD risks such as PI-based regimens. We also observed a significantly increased risk of PTD in a small subset of women on a PI-based regimen compared to women using other more commonly used regimens namely TDF+FTC+EFV. It has previously been hypothesised that the PI ritonavir may directly reduce progesterone levels which could trigger early delivery (103). This could provide a potential mechanistic link between PI-based regimens and the increased risk of PTD we observed.

Fetal growth outcomes (small for gestational age and low birth weight)

As discussed in Chapter 2, optimal placental vasculature growth and function is essential to support *in utero* development. Consequently, we hypothesise that the abnormal fetal growth outcomes we report in this thesis, were a result of disruptions to the placental development process. The placenta is known to be the main source of pro- and anti-angiogenic factors which participate in vasculogenesis and feto-placental angiogenesis (104). Vasculogenic processes in the first trimester, regulate *de novo* formation and growth of blood vessels through remodelling of the uterine spiral arteries (105-107). Angiogenic processes beginning in the second trimester induce remodelling of the underlying placental architecture to allow for increased blood flow and surface area for nutrient exchange (108, 109). Both of these processes mediate placental villi development and maturation, which leads to formation of the materno-fetal interface. An imbalance in pro- and anti-angiogenic factors has been implicated in the occurrence of pregnancy complications such as IUGR (110, 111). Systemic maternal infection and the resultant inflammation has been shown to disrupt the balance of angiogenic factors (112), therefore the consequent placental haemodynamic alterations impact fetal growth and increase the risk of SGA and LBW infants.

While we are unable to distinguish SGA phenotypes (constitutionally small vs growth restricted), we believe that because we observed an increased risk of SGA infants in women living with HIV compared to their -uninfected counterparts (Chapter 5), the SGA we observed was predominantly

due to abnormal pathological processes. The basis for this is that growth restricted SGA infants are most commonly due to maternal factors (HIV infection, hypertension) interacting with placental factors (113). Our findings were regardless of timing of ART initiation, which suggests that different pathways are being triggered by initiation before or during pregnancy.

Our finding of an increased risk of SGA infants among women initiating ART preconception suggests a specific effect on fetal growth during the first trimester through a dysregulation of vasculogenesis. A recent mouse pregnancy model has shown that exposure to preconception ART was associated with excessive small diameter vessel branching in the feto-placental arterial tree that correlate with fetal growth restriction (114). Consequently, we hypothesise that a similar pathway is responsible for women initiating ART preconception being at increased risk of SGA infants, linked to progesterone levels in these women. In addition to inhibiting uterine contractions, progesterone is also a critical regulator of vasculogenesis (107), with lower progesterone levels resulting in disruptions of vasculogenesis (112). Progesterone levels have been shown to be lower in women living with HIV particularly those on PI-based regimens, increasing the risk of adverse birth outcomes related to reduced fetal growth such as SGA (115). Given that in the PIMS cohort only women who initiated ART preconception were the only ones who were on PI-based regimens in the PIMS cohort, this suggests that this proposed pathway likely contributed to the SGA infants seen in this group of women.

In contrast, the increased risk we observed among women initiating during pregnancy was likely linked to dysregulation of angiogenesis. This is because among all women in the PIMS cohort who were Initiating during pregnancy, the majority initiated on ART during the second trimester which coincides with the period of angiogenetic processes. We hypothesise that the ART-mediated immune reconstitution could result in dysregulation of the pathways mediating the placenta being remodelled to support the rapidly growing fetus. Dysregulation these processes has previously been shown to lead to IUGR if the placenta cannot support this rapid fetal growth (112). Further supporting our hypothesis is the observation that maternal HIV infection has been associated with a shift towards an anti-angiogenic state, in the presence of ARV treatment, and this in turn was associated with increased risk of SGA infants (105).

Lastly, our finding of an increased LBW risk linked to abnormal blood pressure trajectories, suggests these findings are attributable to both the effect of early delivery and/or fetal growth restriction (116). While the mechanisms by which maternal blood pressure impact fetal growth are unclear, it has been postulated that it could be linked to placental dysfunction or adverse maternal cardiovascular adaptations to pregnancy (117). These findings as well as our speculation on

potential aetiological pathways points to an urgent need for more mechanistic studies to be conducted alongside epidemiological studies (randomised and observational) providing detailed work on other potential modifiable risk factors, as well as the potential impact that different ART regimens could have on the occurrence of these fetal growth outcomes.

8.2.2.5 Summary

Using methodologically robust data which minimised bias and measurement error, we were able to provide reliable estimates of the risk of adverse birth outcomes associated with HIV and ART status. This thesis presents compelling evidence showing that the risk of PTD in women living with HIV is lower than it previously was, particularly among those using the TDF-FTC-EFV regimen. There also now appears to be little or no increased risk associated with timing of ART initiation. We also provide important evidence pointing to SGA now appearing to be the main problem associated with maternal ART use. The finding that overall there now is a greater risk of SGA infants than of PTD in women living with HIV on ART suggests the importance of drug-specific mechanism occurring at the placental level (19), which warrants further investigation in more focused, mechanistic studies. Findings from this thesis also highlight the importance of efforts to investigate this association using different data sources, through a combination of data from smaller sized prospective studies and larger well-designed population level surveillance databases. Prospective studies will be important for identifying modifiable risk factors and elucidation of aetiological pathways, through the collection of robust and detailed obstetric and safety data. Alongside this, incorporation of population-level surveillance data will be essential to monitor trends in pregnancy outcomes and improve generalisability by including broader populations more reflective of the underlying populations as well as by ensuring sample sizes are sufficiently large to be able to detect clinically relevant differences.

8.3 Strengths and limitations

This thesis has contributed significantly to a greater understanding of the association between maternal ART use and adverse birth outcomes. In addition to the clinical findings, this thesis also contributes to a deeper understanding of the methodological issues that require consideration when designing and interpreting studies investigating not just this association but also other associations involving medication use in pregnancy. However, all findings generated in this thesis need to take into account several important strengths and limitations. While each results chapter included a brief discussion of strengths and limitations related to the specific analyses presented, this section will be a discussion of the overarching strengths and limitations of the body of work presented in this thesis.

The primary strengths of this thesis relate to the robust methodological approaches explored and implemented, with several design features incorporated in order to address some of methodological limitations apparent in previous research of the association of interest. While the majority of analyses used data from the PIMS study, the inclusion of cohort, pregnancy exposure registry and provincial data from the B-Positive study from the same high HIV prevalence setting, enabled confirmation of the results from the PIMS study, as well as a valuable comparison between data sources. The inclusion of women living without HIV from the same communities, who were followed in parallel to their counterparts living with HIV was also a strength, because previous studies investigating this association have been hampered by the absence of an appropriate comparator group women living without HIV. Despite the known concerns around observational studies, the observational nature of the included prospective studies provided good external validity of the experiences of a public sector primary care (low risk) population in pregnancy. The consecutive nature of the two cohort studies, which arose from a single population of pregnant women accessing antenatal care at the GMOU in Cape Town South Africa, enabled a unique opportunity to estimate of the risk of adverse birth outcomes over a longer period of time, while focusing on different aspects (large sample size vs detailed measurement). The PIMS overall cohort enrolled 3952 women with collection of basic cross-sectional and longitudinal clinical data, enabling the detection of trends and patterns of exposures and outcomes. While the B-positive cohort enrolled 989 women and collected additional detailed and robust cross sectional and longitudinal data on psycho-social, economic and behavioural factors which are also risk factors of adverse outcomes, enabling detection of potentially modifiable factors. These sampling and data collection strategies improved the internal validity by minimising confounding. The inclusion of these cohort studies also enabled a valuable direct comparison with previous data for studies investigating the association of interest in the pre-universal ART era from the same facility which enabled us to see the improvements over time which informed interpretation of the findings.

Another strength of this thesis was the use of validated tools with repeated quality assurance for both exposure and outcome assessment, to minimise measurement errors. In the cohort studies included, gestational age was predominantly assessed through ultrasound by a research sonographer using standardised assessment protocols. When ultrasound based measures were not available, gestational age based on SFH measurement was used, which as we determined in Chapter 5 was a reliable alternative method of assessment. This ensured that we were able to minimise the risks of misclassification of PTD and SGA at birth, a necessity when investigating birth outcomes as bias can affect the detection true associations. During the course of the cohort studies, loss to follow-up was minimised by extensive retention measures, while access to routine health care databases enabled outcome ascertainment of women who were lost from face-to-face study follow-up. A further strength of this thesis was the inclusion of routine health care data, especially since the data came from the PHDC, a well-established health information exchange. A beneficial feature of the PHDC is that it also collects valuable clinical data, unlike the administrative claims databases often used to study medication use in pregnancy even though they are created for administrative and billing purposes and have numerous limitations (39). The linkage and de-duplication process of health records employed by the PHDC, as well as the use of a unique provincial patient identifier, strengthened the quality of the data used in this thesis. The comparison of the association of interest across women enrolled in the cohorts' studies, the pregnancy exposure registry (n=9476), and provincial sample (n=183 593), in the same underlying population was illuminating because the similar effect was observed, with the cohort studies and the pregnancy exposure registry providing more details on the associated factors, while the larger province-wide data provided improved precision and enabled the detection of differences between groups. The use of the pregnancy exposure registry and the province wide data not only improved the internal validity of this study by limiting selection bias (whole populations sampled), but also improved external validity of the overall findings because data was available for subsets of women who are typically excluded from prospective studies. Examples of these women include those who did not access antenatal care services during pregnancy and thus were untreated; as well as those with complicated pregnancies which needed specialist care who would have been transferred out of the GMOU, a primary care facility.

The robust methodological approach used in all analyses, where we applied several advanced analytical techniques, such as group-based trajectory modelling, quantitative bias analysis and bias correction of gestational age, is another key strength of this thesis. This allowed us to obtain the least biased effect estimates possible, which would have been unachievable with traditional approaches. Based on prior knowledge of the independent causal relationships between

hypertension, HIV infection and adverse birth outcomes, the blood pressure analysis in this thesis, enabled us to account for this important variable in the main association of interest. We also used exploratory analyses to ensure we identified the best form and expression of blood pressure over the course of pregnancy, and the subsequent approach in which trajectory membership was used as both a dependent (identifying predictors of trajectories) and independent (exploring their impact on future health outcomes) variable provided more accurate insights into the relationships. It should be noted that in its current state GBTM is useful for the type of exploratory analyses conducted in this thesis, however there is limited utility in clinical practice as there is no way to establish the groups at the beginning of follow-up. In recognition of the complex relationships between the primary and secondary exposures, covariates and outcomes, numerous sensitivity analyses were conducted and transparently presented.

Notwithstanding these strengths, there are several key limitations of this thesis which could impact interpretation of our findings and could also contribute to improved study methodology for future investigations of the association between maternal ART and other medications in pregnancy and adverse birth outcomes. The main limitation of the cohort studies was sample size which was insufficient and lead to lack of statistical power, as demonstrated in Chapter 7, to enable us to detect significant, but subtle, differences in PTD risk by HIV status. While the sample sizes were specifically calculated for the PTD outcome measured, particularly in the case of the PIMS study, these calculations were based on much higher expected PTD proportions. This was a combination of the evidence coming from data before the shift to universal test and treat treatment guidelines when delayed ART treatment initiation meant differences by HIV status were much larger; as well as the fact that the majority of the data was from studies using LMP which could have overestimated the PTD proportions.

While the high proportion of women exposed to the WHO recommended TDF+FTC+EFV regimen, particularly among those initiating in pregnancy, is a strength, it also limited the applicability of these findings to settings where women are predominantly receiving the recommended first line DTG-based regimen. Particularly since recent findings from the IMPAACT 2010/VESTED trial showed that DTG+FTC+ TAF was associated with a lower risk of adverse birth outcomes than with TDF-FTC-EFV regimen, so it is possible future studies could find even lower rates of PTD associated with ART use (20). Additionally, the absence of regimen data in the PER and PHDC datasets meant that we were unable to assess the impact of regimens on PTD in the underlying populations where there are likely to be a higher proportion of women who initiated ART preconception using older less safe regimens.

Another limitation is that the observational nature of all the data included in this thesis means that some forms of bias and confounding were still likely. To counteract this we were, as part of the methodological component of this thesis, able to mitigate bias as much as possible in the various analyses. The potential for residual and unmeasured confounding remains likely, despite the multiple measures of potential confounders, with detailed analyses focusing on adjustment and stratification. The lack of precision because of the limited sample size of several subgroups of the stratified analyses also complicated interpretation of results. The reliance of several of the measurement tools on maternal recall, was another limitation, because despite previously being validated, there was still a risk for misclassification.

Lastly, we were unable to further explore any associations in detailed immunological/mechanistic studies. So while we are confident in the quantification of the associations of interest we can only speculate on the reasons for the results we observed.

8.4 Recommendations

As highlighted throughout this thesis, events occurring during pregnancy have important influences on pregnancy outcomes, and in turn newborn health and wellbeing, so understanding the safety of maternal medication use is a public health and moral imperative. Throughout this thesis several recommendations for policy, research practice and future research priorities have been alluded to which could address gaps in the literature of the association of interest. This section summarises these and provides a series of overarching recommendations. Alongside this, additional recommendations will be made related to methodological aspects of perinatal research.

8.4.1 Recommendations for policy

Treatment guidelines are a vital process in the interpretation of research and surveillance evidence and the subsequent transformation into recommendations or policies. Over time more tolerable ARV drugs have been recommended first-line HIV treatment options because evidence has been generated demonstrating their improved safety and efficacy profiles. As discussed in this thesis, pregnant women continue to be excluded from phase I–III studies of investigational agents that do not directly address obstetric conditions (118-120), so evaluation of the safety of new ARVs or treatment options in pregnancy will remain limited to small post-authorisation and cohort studies. While single studies can make valuable contributions to the evidentiary base for treatment guidelines, data from one setting will not provide sufficient evidence of efficacy and/or safety (121). Replication across multiple studies and multiple settings is advantageous and this will only be possible with observational data so HIV treatment guidelines will primarily be based on evidence generated from observational data sources. As highlighted in this Chapter, different regimens have different levels of risk however a main challenge for observational studies in sub-Saharan Africa is the lack of diversity in ART regimens. Consequently, in the future in order to appropriately support guideline changes, head-to-head comparison of specific ART regimens will be required without which it will be difficult to help inform policy.

There is a tendency in some segments of the research community to value novel information over stronger evidence accumulated over time, even though health policies and treatment guidelines require strong evidence. We contend that when evaluating different sources of evidence researchers and policy makers should not only focus on the results, but should also evaluate the methodological qualities of studies by judging the quality of questions, quality of data collection, definition of comparison groups and soundness of analyses. The implications of only focusing on results with no consideration of other methodological components of a study, could have far-reaching consequences as was seen when an early unplanned analysis of the Tsepamo birth outcomes study was conducted. The unexpected identification of a neural tube defect (NTD) safety signal with

preconception DTG caused understandable concern (122). However, despite Botswana's lack of folate supplementation programs (which could have impacted baseline NTD prevalence's), these findings resulted in regulatory reactions that created uncertainty around the most suitable approaches to DTG use among women of reproductive potential (123). Even though subsequent analyses with increased sample size showing the prevalence of NTD's to be lower than in the first analysis, there is still hesitation in some settings to initiate DTG in adolescent girls and women of reproductive age.

The fallout from the Tsepamo study brought into sharp focus the need for reliable data from appropriately designed, studies with excellent data quality and analytical approaches that are planned from study onset (123). Additionally, efforts to strengthen pharmacovigilance around adverse birth outcomes were boosted with recognition of the need for high quality prospective data post-marketing and sentinel site surveillances which can rapidly evaluate safety in pregnancy. While pharmacovigilance systems are well established in high income countries, these are limited in LMICs because there is little budgetary support for these activities from national governments with a heavy reliance on donor support (124, 125). For pharmacovigilance activities to be successful political-will will be necessary to enable establishment of sustainable budgets to develop national pharmacovigilance infrastructure and polices (126). After a number of regional harmonisation efforts (Figure 8.3), the African Medicines Agency has recently been established (127), and it is hoped that this will improve the regulatory and post-marketing surveillance systems across the region. The next steps will need to be significant investment in electronic health information systems enable creation of large healthcare databases capable of using unique individual identifiers for linkage of medication use data with adverse outcomes data (124).

Multisectoral collaboration will be required between key stakeholders including researchers (HIV, maternal and child health and pharmacovigilance), industry and affected communities, to create a common framework for determining the generation and critical evaluation of evidence on the safety of ARVs in pregnancy (123), which will in due course be used to inform policy and clinical care.

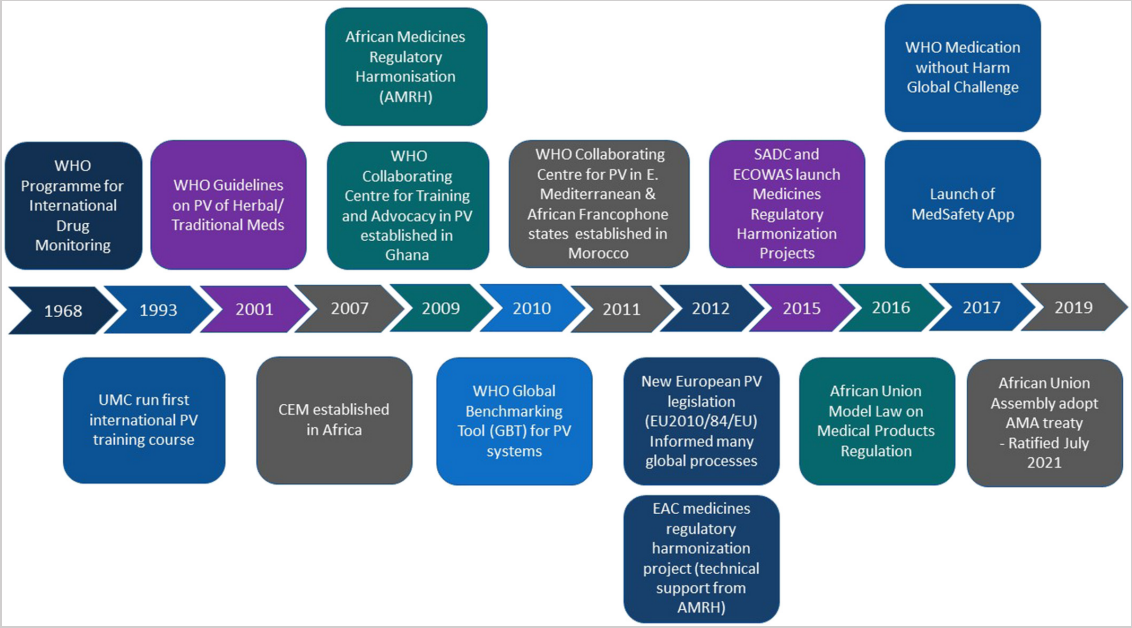

Figure 8.3 Key milestones in the development of pharmacovigilance systems in low- and middle-income countries

Adapted from: Kiguba R, Olsson S, Waitt C. Pharmacovigilance in low-and middle-income countries: A review with particular focus on Africa. *Brit J Clinical Pharma*. 2021; Dec 22.

8.4.2 Recommendations for research practice

Inclusion of pregnant women in clinical trials of medications to be used in pregnancy

Alongside multiple stakeholders, we also call for the inclusion of pregnant women into earlier phases of trials of therapies that will be used by women who are either of reproductive potential or pregnant to avoid their exclusion from receiving new medication or therapy. Despite recent improvements in the regulatory area (128), the current state of research in pregnancy related to medication use has not significantly improved. Even in the face of global pandemics, pregnant women remained excluded from academic-and industry-sponsored investigations critical to the advancement of treatments (129, 130). Therefore these calls apply not only to ART but also other medications for the prevention or treatment of many other serious conditions such as tuberculosis, malaria, and epilepsy (131).

In order to safely and ethically include pregnant women into earlier phases of clinical trials, it will be imperative that these studies are well designed, catering specifically to pregnancy and its unique characteristics as highlighted throughout this thesis. There will also need to be a conceptual shift in how pregnant women are viewed in the context of this type of research (Figure 8.4). Firstly, a change in the designation of pregnant women as being a *"complex population"* instead of as a *"vulnerable population"* (132, 133). Secondly, the notion that women should be protected *"through research"* instead of protected *"from research" should be promulgated* (134). In recent times, rapid advances in SARS-CoV vaccines and therapeutics occurred, however pregnant women had limited

access to these beneficial healthcare interventions because of lack of efficacy, safety and dosing data (135, 136). *Lastly,* as it is a requirement of justice in research*, a shift to* striving for *"fair inclusion"* in research instead of *"presumptive exclusion"* (137).

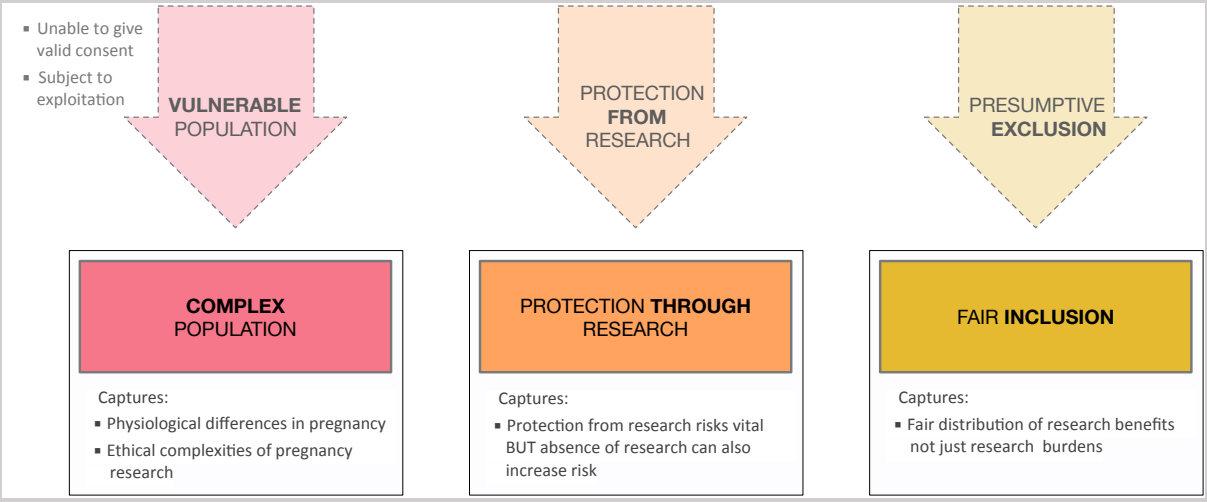

Figure 8.4. Conceptual shift in how pregnant women should be viewed

Actualisation of these conceptual shifts will require future studies to either be opportunistic by keeping women in treatment trials when they become pregnant; or interventional and conducted in pregnant women to identify for the optimal dosages (138). Fast-tracking inclusion rates of women in high prevalence settings, will require *research infrastructure strengthening through global collaborations and multisectoral partnerships.* An example of a successful collaboration between academia and pharmaceutical industry, is the recent Dolphin-2 trial, where a consortium of researchers and pharmaceutical partners in sub-Saharan Africa and Europe, investigated the superiority of DTG-based regimen to EFV-based regimens in women initiating ART late in pregnancy (139, 140).

Improving research harmonisation and reporting guidelines for adverse birth outcomes

Our second recommendation relates to the need for research harmonisation, because of the growing recognition that variability in ascertainment and definitions of exposures and outcomes can undermine the consistent synthesis of the evidence. Lack of consistency in methodology and reporting in research studies investigating adverse birth outcomes, in general, and the association of interest, specifically, has contributed to conflicting findings. This, at times, has been linked to badly designed studies or critical methodological differences which have hampered comparisons across settings and studies (141). This harmonisation can be accomplished through the identification of a core set of variables that should be include in any prospective study on adverse birth outcomes, in

order to improve study consistency independent of primary hypotheses to be evaluated. In addition to outcomes, exposure variables should also be standardised with a need for recognition that studies not considering other key variables associated with adverse birth outcomes, such as blood pressure trajectories and maternal BMI (as highlighted in this thesis), may be short of contextual information required to assess generalisability. The offshoot of this would be improved quality of studies, decreased risk of bias and confounding and the ability to merge numerous datasets that are uniformly structured (141).

As the majority of research investigating the association of interest will utilise observational study design, we would encourage more journals to require the use of the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) guidelines (142), to ascertain high-quality reporting of observational studies. These guidelines aid in ensuring adequate reporting of study procedures (what was planned and what was done), results, assessment of the strengths and weaknesses and conclusions of the study (142, 143). To facilitate the harmonisation proposed, there will need to be the development of more specialised reporting guidelines for prospective observation studies of the outcomes of medication use in pregnancy building on the STROBE guidelines. Similar to the extension of the guidelines for longitudinal observational drug studies in rheumatology (144), this proposed specialised guideline development process would include conducting Delphi surveys, consensus and research team meetings and piloting the checklist. In addition to improving the ability of journal editors, reviewers and readers to critically appraise studies (142), specialised STROBE guidelines will also increase the chances of study's being included in future meta-analyses (28).

8.4.3 Recommendations for future research priorities

This thesis provides good quality observational data from prospective cohorts and routine health care data sources, and also highlights the need for additional observational studies. The priorities of these future studies should focus on assessing the reproducibility of the findings in this thesis, extending generalisability by producing further evidence with other ART regimens or treatment options and providing long-term maternal and infant outcome data following exposure to the new ART regimens. This is particularly important given that the landscape of ART in pregnancy is changing, with the increasing popularity of integrase strand inhibitors (INSTIs) such as DTG and the newer, more potent prodrug of TDF, TAF (145). Additionally, with the use of pre-exposure prophylaxis (PrEP) there will be growing ARV exposure among women living without HIV, which will also need to be monitored (146). Therefore, in addition to smaller prospective studies with detailed, individual-level data, there is also a clear need for ongoing surveillance for ART safety in pregnancy, particularly in LMICs, where the majority of women living with HIV reside.

Another priority will be to have a greater understanding of the changing patterns of ART regimens, patient demographics and risk factors to inform clinical care, intervention development and public health policy (31). This can be facilitated by the improvements in the availability and quality of EHR data and advances in study designs and analytical approaches. These factors have broadened the value of EHR data, by providing a low-cost means of accessing valuable longitudinal data on large populations for surveillance and research purposes (27). The use of larger collaborative studies utilising these routinely collected data sources will be essential since, as highlighted in Chapter 7, population level data will likely now be required to detect differences in adverse birth outcomes, particularly PTD, by HIV status. Another important source of data will be through systematic surveillance for adverse effects, which has essentially been non-existent in LMICs because of poor medication exposure in pregnancy documentation, high proportion of deliveries outside the health care setting, and the lack of sufficient resources to prospectively follow women and their children when little to no research infrastructure exists (145). Consequently, more investment will be required into systems such as the PHDC in the Western Cape, which are well designed and maintained, to ensure high quality population level surveillance data, which with appropriate linkage could potentially allow for thelong-term follow up of exposed children.

There is an unmet clinical need to improve accurate prediction of blood pressure-related adverse outcomes early in pregnancy. This could be improved clinically and epidemiologically through two lines of research and clinician application which have shown promise and will need to be prioritised in future research work. Clinically, the traditional care pathway can be transformed through remote monitoring of the blood pressure, clinical symptoms, and urine protein levels in high-risk pregnant women (147). Epidemiologically, trajectory modelling enables the identification of subgroups of individuals with similar trajectory patterns which can be useful to inform prevention and clinical practice (148). To improve this modelling approaches' utility for prediction of pregnancy related conditions, several updates could be made to the current group based trajectory models. The most natural approach will be to update the trajectory modelling as a function of gestational age either in a time dependent covariates model or through use a of a causal inference approach. Another approach could be the incorporation of gestational age specific blood pressure centiles into the model. Predictive models for complex conditions are also improving through the adoption of machine learning approaches. While there has been limited evidence generated using machine learning to predict the risk of blood pressure related outcomes in pregnancy these techniques have great potential to improve the diagnostic process and treatment of preeclampsia and possibly other obstetric conditions (149). These machine learning approaches are often superior to "traditional" analytic approaches because they are able to infer complex, highly non-linear relations between

statistical inputs (147). Consequently future research efforts should seek to utilise these approaches in order to increase the evidence base around their use to ultimately improve care in pregnancy.

This thesis has also highlighted the need for a clearer understanding of immunological pathways involved in normal pregnancy. The generalisation of pregnancy as a condition of immune suppression has created a myth of pregnancy as a state of immunological weakness and therefore of increased susceptibility to infectious diseases (150). Consequently, it will be important to tease out how specific endocrinological, physiological and immunological factors increase the risk of infection and in turn adverse birth outcomes will require careful considerations (151). Additionally, further immunological research will be required into the relationship between maternal HIV and ART use and the cytokine environment during pregnancy, in order to decipher the biological pathway between ART exposure, and SGA infants. The immunological mechanisms underlying other pregnancy complications in women living with HIV, such as pre-eclampsia and stillbirth, also need to be determined.

There is currently limited data on placental development processes, in the context of HIV infection, which as discussed in this Chapter could lead to persistent dysregulation of inflammatory and angiogenic pathways. Given that these pathways are critical to placental vascular development, HIV/ART mediated disruptions of these process could potentially provide an explanation for why, even in the face of appropriate antiretroviral therapy, rates of adverse birth outcomes remain high more placenta studies will be required, especially in settings where the highest rates adverse birth outcomes occur (114). This suggests that alternative strategies may be required for infections in pregnancy, with a shift from those that only target the pathogen, to those that focus on modifying host response pathways that may be the driving force behind the occurrence of adverse birth outcomes (112).

8.5 Conclusions

In this thesis exploring the association between maternal ART use and adverse birth outcomes in Cape Town, South Africa, we applied a methodological robust approach in order to minimise the biases, where possible, that likely affected previous studies and impacted interpretation of findings. Using bias-corrected gestational age, we found that the previously observed excess risks of PTD among women living with HIV compared to their counterparts living without HIV now appear to be reduced, however this may not be fully ameliorated, under policies of universal maternal ART use. Another important finding was the increased risk of SGA infants among women living with HIV on ART, which is particularly concerning given that these infants are at significantly increased risk of short- and long-term health outcomes. These findings point to the need for additional research on

the immunological pathways through which maternal HIV infection and ART exposure influence the fetal environment, with the view to design treatment options that may benefit both mother and fetus, while minimising risks.

The benefits of ART for maternal and infant health are indisputable. However, given continued questions around the risk of adverse birth outcomes, particularly with newer ART regimens, future research and surveillance will need to evaluate the comparative safety of specific ART regimens alongside the strengthening of the systems to assess medication safety. This thesis is an important packaging of reproductive and perinatal issues which tend to only be focused on by perinatal epidemiologists, which need to be more widely considered and applied when designing studies or routine surveillance systems investigating medications use in pregnancy.

8.6 References

1. Abrams EJ, Myer L. Can we achieve an AIDS-free generation? Perspectives on the global campaign to eliminate new pediatric HIV infections. *J Acquir Immune Defic Syndr*. 2013;63(Suppl 2):S208-S212.
2. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: summary of key features and recommendations. Geneva, Switzerland: World Health Organization; 2013.
3. UNAIDS. Seizing the moment — Tackling entrenched inequalities to end epidemics. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2020.
4. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. Geneva, Switzerland: World Health Organization; 2016.
5. Goga AE, Dinh TH, Essajee S, et al. What will it take for the Global Plan priority countries in Sub-Saharan Africa to eliminate mother-to-child transmission of HIV? *BMC Infect Dis*. 2019;19(Suppl 1):783-796.
6. WHO. Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV: interim guidelines: supplement to the 2016 consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: World Health Organization, 2018.
7. UNAIDS. Factsheet: global AIDS update. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2021.
8. Zash R, Souda S, Chen JY, et al. Reassuring birth outcomes with tenofovir/emtricitabine/efavirenz used for prevention of mother-to-child transmission of HIV in Botswana. *J Acquir Immune Defic Syndr*. 2016;71(4):428-436.
9. Bengtson AM, Chibwesha CJ, Westreich D, et al. Duration of cART before delivery and low infant birthweight among HIV-infected women in Lusaka, Zambia. *J Acquir Immune Defic Syndr*. 2016;71(5):563-569.
10. Boer K, Nellen JF, Patel D, et al. The AmRo study: pregnancy outcome in HIV-1-infected women under effective highly active antiretroviral therapy and a policy of vaginal delivery. *Br J Obstet Gynaecol.* 2007;114(2):148-155.
11. Kesho Bora Study Group, de Vincenzi I. Triple antiretroviral compared with zidovudine and single-dose nevirapine prophylaxis during pregnancy and breastfeeding for prevention of mother-to-child transmission of HIV-1 (Kesho Bora study): a randomised controlled trial. *Lancet Infect Dis*. 2011;11(3):171-180.

12. Njom Nlend AE, Nga Motazé A, Moyo Tetang S, et al. Preterm Birth and Low Birth Weight after In Utero Exposure to Antiretrovirals Initiated during Pregnancy in Yaoundé, Cameroon. *PLoS One*. 2016;11(3):e0150565.
13. Olagbuji BN, Ezeanochie MC, Ande AB, Oboro VO. Obstetric and perinatal outcome in HIV positive women receiving HAART in urban Nigeria. *Arch Gynecol Obstet*. 2010;281(6):991-994.
14. Tuomala RE, Watts DH, Li D, et al. Improved obstetric outcomes and few maternal toxicities are associated with antiretroviral therapy, including highly active antiretroviral therapy during pregnancy. *J Acquir Immune Defic Syndr*. 2005;38(4):449-473.
15. Traisathit P, Mary JY, Le Cœur S, et al. Risk factors of preterm delivery in HIV-infected pregnant women receiving zidovudine for the prevention of perinatal HIV. *J Obstet Gynaecol Res*. 2009 Apr;35(2):225-233.
16. Young S, Murray K, Mwesigwa J, et al. Maternal nutritional status predicts adverse birth outcomes among HIV-infected rural Ugandan women receiving combination antiretroviral therapy. *PLoS One*. 2012;7(8):e41934.
17. Liu KC, Farahani M, Mashamba T, et al. Pregnancy outcomes and birth defects from an antiretroviral drug safety study of women in South Africa and Zambia. *AIDS.* 2014;28(15):2259-2268.
18. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.
19. Zash R, Jacobson DL, Diseko M, et al. Comparative safety of antiretroviral treatment regimens in pregnancy. *JAMA Pediatr*. 2017;171(10):e172222-e
20. Lockman S, Brummel SS, Ziemba L, et al. Efficacy and safety of dolutegravir with emtricitabine and tenofovir alafenamide fumarate or tenofovir disoproxil fumarate, and efavirenz, emtricitabine, and tenofovir disoproxil fumarate HIV antiretroviral therapy regimens started in pregnancy (IMPAACT 2010/VESTED): a multicentre, open-label, randomised, controlled, phase 3 trial. *Lancet.* 2021;397(10281):1276-1292.
21. Paneth N, Joyner M. The use of observational research to inform clinical practice. The Journal of Clinical Investigation. 2021;131(2).
22. Tshivuila-Matala CO, Honeyman S, Nesbitt C, et al. Adverse perinatal outcomes associated with antiretroviral therapy regimens: systematic review and network meta-analysis. *AIDS*. 2020;34(11):1643-1656.
23. Satija A, Yu E, Willett WC, Hu FB. Understanding nutritional epidemiology and its role in policy. *Adv Nutr*. 2015;6(1):5-18.
24. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342(25):1878-1886.

25. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887-1892.
26. Canova C, Cantarutti A. Population-Based Birth Cohort Studies in Epidemiology. Multidisciplinary Digital Publishing Institute; 2020. p. 5276.
27. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health*. 2016;37:61-81.
28. Egger M, Schneider M, Smith GD. Meta-analysis Spurious precision? Meta-analysis of observational studies. *BMJ*. 1998;316(7125):140-144.
29. Haneef R, Delnord M, Vernay M, et al. Innovative use of data sources: a cross-sectional study of data linkage and artificial intelligence practices across European countries. *Arch Public Health*. 2020;78(1):1-11.
30. Ferrante A. The use of data-linkage methods in criminal justice research: a commentary on progress, problems and future possibilities. *Curr Issues Crim Justice*. 2009;20(3):378-392.
31. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. *Diabetologia*. 2018;61(6):1241-1248.
32. Johnson LF, Mutemaringa T, Heekes A, Boulle A. Effect of HIV Infection and Antiretroviral Treatment on Pregnancy Rates in the Western Cape Province of South Africa. *J Infect Dis.* 2020;221(12):1953-1962.
33. UNAIDS. AIDS targets 2025. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS. https://aidstargets2025.unaids.org (accessed on 23 June 2022)
34. 31. UNAIDS. Understanding fast-track: Accelerating action to end the AIDS epidemic by 2030. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2015.
35. Mehta U, Heekes A, Kalk E, Boulle A. Assessing the value of Western Cape Provincial Government health administrative data and electronic pharmacy records in ascertaining medicine use during pregnancy. *S Afr Med J*. 2018;108(5):439-443.
36. Artama M, Gissler M, Malm H, Ritvanen A, Drugs and Pregnancy Study Group. Nationwide register-based surveillance system on drugs and pregnancy in Finland 1996–2006. *Pharmacoepidemiol Drug Saf*. 2011;20(7):729-738.
37. Howard TB, Tassinari MS, Feibus KB, Mathis LL, editors. Monitoring for teratogenic signals: pregnancy registries and surveillance methods. American Journal of Medical Genetics Part C: Seminars in Medical Genetics; 2011: Wiley Online Library.
38. Andrade SE, Raebel MA, Morse AN, et al. Use of prescription medications with a potential for fetal harm among pregnant women. *Pharmacoepidemiol Drug Saf*. 2006;15(8):546-554.

39. Andrade SE, Bérard A, Nordeng HM, et al. Administrative claims data versus augmented pregnancy data for the study of pharmaceutical treatments in pregnancy. *Curr Epidemiol Rep*. 2017;4(2):106-116.
40. Charlton RA, Neville AJ, Jordan S, et al. Healthcare databases in Europe for studying medicine use and safety during pregnancy. *Pharmacoepidemiol Drug Saf*. 2014;23(6):586-594.
41. Kalk E, Metha U, Slogrove AL, et al. Pregnancy Exposure Registry/Birth Defects Surveillance Programme In The Western Cape, South Africa: A Model For Low- & Middle-Income Countries. 18th Annual meeting International Society of Pharmacovigilance, Geneva, Switzerland, 2018.
42. Rothman KJ. Induction and latent periods. *Am J Epidemiol.* 1981;114(2):253-259.
43. Wood ME, Lupattelli A, Palmsten K, et al. Longitudinal methods for modeling exposures in pharmacoepidemiologic studies in pregnancy. *Epidemiol Rev*. 2021;43(1):130-146.
44. Lee TA, Pickard AS. Exposure definition and measurement. Developing a protocol for observational comparative effectiveness research: a user's guide: Agency for Healthcare Research and Quality (US); 2013.
45. Woldesenbet S, Kufa T, Cheyip M, et al. Awareness of HIV-positive status and linkage to treatment prior to pregnancy in the “test and treat” era: A national antenatal sentinel survey, 2017, South Africa. *PLoS One*. 2020;15(3):e0229874.
46. de Beer S, Kalk E, Kroon M, Boulle A, Osler M, Euvrard J, et al. A longitudinal analysis of the completeness of maternal HIV testing, including repeat testing in Cape Town, South Africa. *J Int AIDS Soc*. 2020;23(1):e25441.
47. Grzeskowiak LE, Gilbert AL, Morrison JL. Exposed or not exposed? Exploring exposure classification in studies using administrative data to investigate outcomes following medication use during pregnancy. *Eur J Clin Pharmacol*. 2012;68(5):459-467.
48. Källén B, Nilsson E, Olausson PO. Antidepressant use during pregnancy: comparison of data obtained from a prescription register and from antenatal care records. *European J Clin Pharmacol*. 2011;67(8):839-845.
49. van der Hoven J, Allen E, Cois A, et al. Determining antenatal medicine exposures in South African women: a comparison of three methods of ascertainment. *BMC Pregnancy Childbirth*. 2022;22(1):1-15.
50. Velentgas P, Dreyer NA, Wu AW. Outcome definition and measurement. Developing a protocol for observational comparative effectiveness research: a user's guide: Agency for Healthcare Research and Quality (US); 2013.
51. Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic measures of the course and outcome of pregnancy. *Epidemiol Rev*. 2002;24(2):91-101.
52. Jackson D, Lang J, Ganiats T. Epidemiological issues in perinatal outcomes research. *Paediatr Perinat Epidemiol*. 1999;13(4):392-404.

53. Villar J, Papageorghiou AT, Knight HE, et al. The preterm birth syndrome: a prototype phenotypic classification. *Am J Obstet Gynecol.* 2012;206(2):119-123.
54. Manuck TA, Esplin MS, Biggio J, et al. The phenotype of spontaneous preterm birth: application of a clinical phenotyping tool. *Am J Obstet Gynecol*. 2015;212(4):487.e1-e11
55. Olsen J, Basso O. Reproductive epidemiology. In: Ahrens W PI (Ed) Handbook of Epidemiology. pg 1043–1090. Springer Verlag: Berlin, Germany; 2005.
56. Messerlian C, Basso O. Cohort studies in the context of obstetric and gynecologic research: a methodologic overview. *Acta Obstet Gynecol Scand*. 2018;97(4):371-379.
57. Arbogast PG, VanderWeele TJ. Considerations for statistical analysis. Developing a protocol for observational comparative effectiveness research: a user's guide: Agency for Healthcare Research and Quality (US); 2013.
58. Pazzagli L, Linder M, Zhang M, et al. Methods for time-varying exposure related problems in pharmacoepidemiology: An overview. *Pharmacoepidemiol Drug Saf*. 2018;27(2):148-160.
59. Neophytou AM, Kioumourtzoglou MA, Goin DE, Darwin KC, Casey JA. Educational note: addressing special cases of bias that frequently occur in perinatal epidemiology. *Int J Epidemiol.* 2021;50(1):337-345.
60. Gasparrini A. Modeling exposure–lag–response associations with distributed lag non-linear models. *Stat Med*. 2014;33(5):881-899.
61. Wagner M, Grodstein F, Leffondre K, Samieri C, Proust-Lima C. Time-varying associations between an exposure history and a subsequent health outcome: a landmark approach to identify critical windows. *BMC Med Res Methodol*. 2021;21(1):1-15.
62. Hermida RC, Ayala DE, Iglesias M. Predictable blood pressure variability in healthy and complicated pregnancies. *Hypertension*. 2001;38(3):736-741.
63. Maddox TM, Ross C, Tavel HM, et al. Blood pressure trajectories and associations with treatment intensification, medication adherence, and outcomes among newly diagnosed coronary artery disease patients. *Circ Cardiovasc Qual Outcomes*. 2010;3(4):347-357.
64. Prada-Ramallal G, Takkouche B, Figueiras A. Bias in pharmacoepidemiologic studies using secondary health care databases: a scoping review. *BMC Med Res Methodol*. 2019;19(1):1-14.
65. Wood ME, Lapane KL, van Gelder MM, Rai D, Nordeng HM. Making fair comparisons in pregnancy medication safety studies: An overview of advanced methods for confounding control. *Pharmacoepidemiol Drug Saf*. 2018;27(2):140-147.
66. Wedi CO, Kirtley S, Hopewell S, et al. Perinatal outcomes associated with maternal HIV infection: a systematic review and meta-analysis. *Lancet HIV*. 2016;3(1):e33-e48.
67. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.

68. Kourtis AP, Schmid CH, Jamieson DJ, Lau J. Use of antiretroviral therapy in pregnant HIV-infected women and the risk of premature delivery: a meta-analysis. *AIDS*. 2007;21(5):607-615.
69. Watts DH, Williams PL, Kacanek D, et al. Combination antiretroviral use and preterm birth. *J Infect Dis*. 2013;207(4):612-621.
70. Sibiude J, Warszawski J, Tubiana R, et al. Premature delivery in HIV-infected women starting protease inhibitor therapy during pregnancy: role of the ritonavir boost? *Clin Infect Dis.* 2012;54(9):1348-1360.
71. Ezechi O, Oladele D, Kalejaiye O, Oke B, Ohwodo H, Adu R, et al. Pregnancy, obstetric and neonatal outcomes in HIV positive Nigerian women. *Afr J Reprod Health*. 2013;17(3):160-168.
72. Cooke R. Catch-up growth: implications for the preterm and term infant. *Eur J Clin Nutr*. 2010;64(Suppl 1):S8-S10.
73. Siemieniuk RA, Lytvyn L, Ming JM, et al. Antiretroviral therapy in pregnant women living with HIV: a clinical practice guideline. *BMJ*. 2017;358: j3961
74. Romo ML, Patel RC, Edwards JK, et al. Disparities in Dolutegravir Uptake Affecting Females of Reproductive Age With HIV in Low-and Middle-Income Countries After Initial Concerns About Teratogenicity: An Observational Study. *Ann Intern Med*. 2022;175(1):84-94.
75. Kanters S, Vitoria M, Zoratti M, et al. Comparative efficacy, tolerability and safety of dolutegravir and efavirenz 400mg among antiretroviral therapies for first-line HIV treatment: a systematic literature review and network meta-analysis. *EClinicalMedicine*. 2020;28:100573.
76. Phillips AN, Venter F, Havlir D, et al. Risks and benefits of dolutegravir-based antiretroviral drug regimens in sub-Saharan Africa: a modelling study. *Lancet HIV*. 2019;6(2):e116-e27.
77. Svensson AC, Pawitan Y, Cnattingius S, Reilly M, Lichtenstein P. Familial aggregation of small-for-gestational-age births: the importance of fetal genetic effects. *Am J Obstet Gynecol.* 2006;194(2):475-479.
78. Kok JH, Lya den Ouden A, Verloove-Vanhorick SP, Brand R. Outcome of very preterm small for gestational age infants: the first nine years of life. *Br J Obstet Gynaecol.* 1998;105(2):162-168.
79. Griffin I, Lee H, Profit J, Tancedi D. The smallest of the small: short-term outcomes of profoundly growth restricted and profoundly low birth weight preterm infants. *J Perinatol.* 2015;35(7):503-510.
80. Baer R, Rogers E, Partridge J, et al. Population-based risks of mortality and preterm morbidity by gestational age and birth weight. *J Perinatol*. 2016;36(11):1008-1013.
81. Adair LS, Guilkey DK. Age-specific determinants of stunting in Filipino children. *J Nutr.* 1997;127(2):314-320.
82. Zhang J, Merialdi M, Platt LD, Kramer MS. Defining normal and abnormal fetal growth: promises and challenges. *Am J Obstet Gynecol*. 2010;202(6):522-528.

83. Weiner C, Sabbagha R, Vaisrub N, Depp R. A hypothetical model suggesting suboptimal intrauterine growth in infants delivered preterm. *Obstet Gynecol*. 1985;65(3):323-326.
84. Zeitlin J, Ancel P, Saurel-Cubizolles M, Papiernik E. The relationship between intrauterine growth restriction and preterm delivery: an empirical approach using data from a European case-control study. *Br J Obstet Gynaecol*. 2000;107(6):750-758.
85. Morken N-H, Källen K, Jacobsson B. Fetal growth and onset of delivery: a nationwide population-based study of preterm infants. *Am J Obstet Gynecol*. 2006;195(1):154-161.
86. Sykes L, MacIntyre DA, Yap XJ, Teoh TG, Bennett PR. The Th1:Th2 Dichotomy of Pregnancy and Preterm Labour. *Mediators Inflamm*. 2012;2012:967629.
87. Fiore S, Ferrazzi E, Newell ML, Trabattoni D, Clerici M. Protease Inhibitor—Associated Increased Risk of Preterm Delivery Is an Immunological Complication of Therapy*. J Infect Dis.* 2007;195(6):914-916.
88. Berger A. Th1 and Th2 responses: what are they? *BMJ*. 2000;321(7258):424.
89. Norman JE. Progesterone and preterm birth. *Int J Gynaecol Obstet*. 2020;150(1):24-30.
90. Vyas P, Mathad JS, Leu CS, et al. Impact of HIV status on systemic inflammation during pregnancy. *AIDS*. 2021;35(14):2259-2268.
91. Khong TY, Mooney EE, Ariel I, et al. Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement. *Arch Pathol Lab Med*. 2016;140(7):698-713.
92. Ernst LM. Maternal vascular malperfusion of the placental bed. *APMIS*. 2018;126(7):551-560.
93. Schwartz DA, Sungkarat S, Shaffer N, et al. Placental abnormalities associated with human immunodeficiency virus type 1 infection and perinatal transmission in Bangkok, Thailand. *J Infect Dis.* 2000;182(6):1652-1657.
94. Obimbo MM, Zhou Y, McMaster MT, et al. Placental structure in preterm birth among HIV-positive versus HIV-negative women in Kenya. *J Acquir Immune Defic Syndr*. 2019;80(1):94-102.
95. Smid MC, Stringer EM, Stringer JS. A worldwide epidemic: the problem and challenges of preterm birth in low-and middle-income countries. *Am J Perinatol*. 2016;33(03):276-289.
96. Kramer MS, Papageorghiou A, Culhane J, et al. Challenges in defining and classifying the preterm birth syndrome. *Am J Obstet Gynecol.* 2012;206(2):108-112.
97. Fiore S, Newell ML, Trabattoni D, et al. Antiretroviral therapy-associated modulation of Th1 and Th2 immune responses in HIV-infected pregnant women. *J Reprod Immunol*. 2006;70(1):143-150.
98. Ohno Y, Kasugai M, Kurauchi O, Mizutani S, Tomoda Y. Effect of interleukin 2 on the production of progesterone and prostaglandin E2 in human fetal membranes and its consequences for preterm uterine contractions. *Eur J Endocrinol*. 1994;130(5):478-484.

99. Powis KM, Shapiro RL. Protease inhibitors and adverse birth outcomes: is progesterone the missing piece to the puzzle? *J Infect Dis*. 2014;211(1):4-7.

100. Stoner MC, Cole SR, Price J, Winston J, Stringer JS. Timing of initiation of antiretroviral therapy and risk of preterm birth in studies of HIV-infected pregnant women: The role of selection bias. *Epidemiology*. 2018;29(2):224-229.

101. Stringer JS, Stoner MC, Kasaro MP, Vwalika B, Cole SR. Preconception ART and preterm birth: real effect or selection bias. *Lancet HIV*. 2017;4(4):e150.

102. Ikumi NM, Malaba TR, Pillay K, et al. Differential impact of antiretroviral therapy initiated before or during pregnancy on placenta pathology in HIV-positive women. *AIDS.* 2021;35(5):717-726.

103. Rough K, Seage GR 3rd, Williams PL, et al. Birth Outcomes for Pregnant Women with HIV Using Tenofovir-Emtricitabine. *N Engl J Med*. 2018;378(17):1593-1603.

104. Djeha A, Girard S, Trottier H, et al. No association between early antiretroviral therapy during pregnancy and plasma levels of angiogenic factors: a cohort study. *BMC Pregnancy Childbirth*. 2019;19(1):1-10.

105. Conroy AL, McDonald CR, Gamble JL, et al. Altered angiogenesis as a common mechanism underlying preterm birth, small for gestational age, and stillbirth in women living with HIV. *Am J Obstet Gynecol*. 2017;217(6):684-701.

106. Kim M, Park HJ, Seol JW, et al. VEGF-A regulated by progesterone governs uterine angiogenesis and vascular remodelling during pregnancy. *EMBO Mol Med*. 2013;5(9):1415-1430.

107. Chen JZ, Wong MH, Brennecke SP, Keogh RJ. The effects of human chorionic gonadotrophin, progesterone and oestradiol on trophoblast function. *Mol Cell Endocrinol*. 2011;342(1-2):73-80.

108. Geva E, Ginzinger DG, Zaloudek CJ, et al. Human placental vascular development: vasculogenic and angiogenic (branching and nonbranching) transformation is regulated by vascular endothelial growth factor-A, angiopoietin-1, and angiopoietin-2. *J Clin Endocrinol Metab*. 2002;87(9):4213-4224.

109. Charnock-Jones D, Kaufmann P, Mayhew TM. Aspects of human fetoplacental vasculogenesis and angiogenesis. 1. Molecular regulation. *Placenta*. 2004;25(2-3):103-113.

110. Romero R, Nien JK, Espinoza J, et al. A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. *J Matern Fetal Neonatal Med*. 2008;21(1):9-23.

111. Taylor RN, Grimwood J, Taylor RS, et al. Longitudinal serum concentrations of placental growth factor: evidence for abnormal placental angiogenesis in pathologic pregnancies. *Am J Obstet Gynecol*. 2003;188(1):177-182.
112. Weckman AM, Ngai M, Wright J, McDonald CR, Kain KC. The impact of infection in pregnancy on placental vascular development and adverse birth outcomes. *Front Microbiol*. 2019;10(1924):1-11.
113. Osuchukwu OO, Reed DJ. Small for Gestational Age. In: StatPearls [Internet]. Florida, USA: StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK563247/ (accessed 15 January 2022)
114. Mohammadi H, Papp E, Cahill L, et al. HIV antiretroviral exposure in pregnancy induces detrimental placenta vascular changes that are rescued by progesterone supplementation. *Sci Rep*. 2018;8(1):1-14.
115. Papp E, Balogun K, Banko N, et al. Low prolactin and high 20-α-hydroxysteroid dehydrogenase levels contribute to lower progesterone levels in HIV-infected pregnant women exposed to protease inhibitor–based combination antiretroviral therapy. *J Infect Dis.* 2016;213(10):1532-1540.
116. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull WHO.* 1987;65:663–737.
117. Bakker R, Steegers EA, Hofman A, Jaddoe VW. Blood pressure in different gestational trimesters, fetal growth, and the risk of adverse birth outcomes: the generation R study. *Am J Epidemiol.* 2011;174(7):797-806.

118. Andrade SE, Gurwitz JH, Davis RL, et al. Prescription drug use in pregnancy. *Am J Obstet Gynecol.* 2004;191(2):398-407.
119. Chambers CD, Polifka JE, Friedman JM. Drug safety in pregnant women and their babies: ignorance not bliss. *Clin Pharmacol Ther*. 2008;83(1):181-183.
120. Kaye DK. The moral imperative to approve pregnant women's participation in randomized clinical trials for pregnancy and newborn complications. *Philos Ethics Humanit Med*. 2019;14(1):1-11.
121. Hollon D, Miller IJ, Robinson E. Criteria for evaluating treatment guidelines. *Am Psychol*. 2002;57(12):1052-1059.
122. Zash R, Makhema J, Shapiro RL. Neural-tube defects with dolutegravir treatment from the time of conception. *N Engl J Med.* 2018;379(10):979-981.
123. Mofenson LM, Pozniak AL, Wambui J, et al. Optimizing responses to drug safety signals in pregnancy: the example of dolutegravir and neural tube defects. *J Int AIDS Soc.* 2019;22(7):e25352.

124. Kiguba R, Olsson S, Waitt C. Pharmacovigilance in low-and middle-income countries: A review with particular focus on Africa. *Br J Clin Pharmacol*. 2021;1-19
125. Ampadu HH, Hoekman J, Arhinful D, et al. Organizational capacities of national pharmacovigilance centres in Africa: assessment of resource elements associated with successful and unsuccessful pharmacovigilance experiences. *Global Health.* 2018;14(1):109-126.
126. Olsson S, Pal SN, Dodoo A. Pharmacovigilance in resource-limited countries. *Expert Rev Clin Pharmacol*. 2015;8(4):449-460.
127. Ncube BM, Dube A, Ward K. Establishment of the African Medicines Agency: progress, challenges and regulatory readiness. *J Pharm Policy Pract*. 2021;14(1):29-41
128. Ballantyne A. Pregnant women can finally expect better. *Hastings Cent Rep*. 2019;49(1):10-11.
129. Gomes MF, de la Fuente-Núñez V, Saxena A, Kuesel AC. Protected to death: systematic exclusion of pregnant women from Ebola virus disease trials. *Reprod Health*. 2017;14(Suppl 3):47-55.
130. Smith DD, Pippen JL, Adesomo AA, et al. Exclusion of pregnant women from clinical trials during the coronavirus disease 2019 pandemic: a review of international registries. *Am J Perinatol.* 2020;37(8):792-799.
131. Gupta A, Mathad JS, Abdel-Rahman SM, et al. Toward earlier inclusion of pregnant and postpartum women in tuberculosis drug trials: consensus statements from an international expert panel. *Clin Infect Dis*. 2016;62(6):761-769.
132. ACOG. Committee opinion no. 646: ethical considerations for including women as research participants. *Obstet Gynecol*. 2015;126:e100-7.
133. Krubiner CB, Faden RR. Pregnant women should not be categorised as a 'vulnerable population'in biomedical research studies: ending a vicious cycle of 'vulnerability'. *J Med Ethics*. 2017;43(10):664-665.
134. Lyerly AD, Beigi R, Bekker LG, et al. Ending the evidence gap for pregnancy, HIV and co-infections: ethics guidance from the PHASES project. *J Int AIDS Soc.* 2021;24(12):e25846.
135. Beigi RH, Krubiner C, Jamieson DJ, et al. The need for inclusion of pregnant women in COVID-19 vaccine trials. *Vaccine.* 2021;39(6):868.
136. Taylor MM, Kobeissi L, Kim C, Amin A, Thorson AE, Bellare NB, et al. Inclusion of pregnant women in COVID-19 treatment trials: a review and global call to action. *Lancet Glob Health*. 2021;9(3):e366-e371.
137. MacKay D, Saylor KW. Four faces of fair subject selection. *Am J Bioeth.* 2020;20(2):5-19.
138. Colbers A, Mirochnick M, Schalkwijk S, et al. Importance of prospective studies in pregnant and breastfeeding women living with human immunodeficiency virus. *Clin Infect Dis*. 2019;69(7):1254-1258.

139. Waitt C, Orrell C, Walimbwa S, et al. Safety and pharmacokinetics of dolutegravir in pregnant mothers with HIV infection and their neonates: a randomised trial (DolPHIN-1 study). *PLoS Med*. 2019;16(9):e1002895.

140. Kintu K, Malaba TR, Nakibuka J, et al. Dolutegravir versus efavirenz in women starting HIV therapy in late pregnancy (DolPHIN-2): an open-label, randomised controlled trial. *Lancet HIV*. 2020;7(5):e332-e9.

141. Khalil A, Gordijn SJ, Beune IM, et al. Essential variables for reporting research studies on fetal growth restriction: a Delphi consensus. *Ultrasound Obstet Gynecol*. 2019;53(5):609-614.

142. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med.* 2007;147(8):W163-94.

143. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007;147(8):573-577.

144. Zavada J, Dixon WG, Askling J. Launch of a checklist for reporting longitudinal observational drug studies in rheumatology: a EULAR extension of STROBE guidelines based on experience from biologics registries. *Ann Rheum Dis.* 2014;73(3):628.

145. Zash RM, Williams PL, Sibiude J, Lyall H, Kakkar F. Surveillance monitoring for safety of in utero antiretroviral therapy exposures: current strategies and challenges. *Expert Opin Drug Saf*. 2016;15(11):1501-1513.

146. Heffron R, Pintye J, Matthews LT, Weber S, Mugo N. PrEP as peri-conception HIV prevention for women and men. *Curr HIV/AIDS Rep*. 2016;13(3):131-139.

147. Hackelöer M, Schmidt L, Verlohren S. New advances in prediction and surveillance of preeclampsia: role of machine learning approaches and remote monitoring. *Arch Gynecol and Obstet*. 2022:1-15.

148. Nguena Nguefack HL, Pagé MG, Katz J, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. *Clin Epidemiol.* 2020;12:1205-1222.

149. Lee SM, Nam Y, Choi ES, et al. Development of early prediction model for pregnancy-associated hypertension with graph-based semi-supervised learning. *Sci Rep*. 2022;12(1):15793.

150. Mor G, Cardenas I. The immune system in pregnancy: a unique complexity. *Am J Reprod Immunol*. 2010 Jun;63(6):425-433.

151. Abu-Raya B, Michalski C, Sadarangani M, Lavoie PM. Maternal Immunological Adaptation During Normal Pregnancy. *Front Immunol*. 2020;11:575197.

Chapter 9: Appendices

9.1 Ethical approvals

9.1.1 University of Cape Town Human Research Ethics Committee for the PIMS study

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
Email: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

30 October 2014

HREC REF: 739/2014

A/Prof L Myer
Epidemiology & Biostatistics
Public Health & Family Medicine
Falmouth Building
Level 5, entrance 5

Dear A/Prof Myer

PROJECT TITLE: ANTIRETROVIRAL THERAPY AND RISK OF PREMATURE DELIVERY: THE PREMATURITY IMMUNOLOGY IN HIV-INFECTED MOTHERS AND THEIR INFANTS STUDY (PIMS)

Thank you for your response to the Faculty of Health Sciences Human Research Ethics Committee dated 28 October 2014.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30th October 2015.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Yours sincerely

Signed by candidate

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.

HREC 739/2014

HUMAN RESEARCH
ETHICS COMMITTEE
24 MAY 2022
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.5.23
☐ Not approved	See attached comments		
Signature Chairperson of the HREC	Signed by candidate	Date Signed	24/5/2022

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	20th May 2022		
HREC REF Number	739/2014	Current Ethics Approval was granted until	30/03/2022
Protocol title	Antiretroviral treatment during pregnancy: understanding the risk of premature delivery: The Prematurity Immunology in HIV-infected Mothers and their infants Study (PIMS)		
Protocol number (if applicable)	N/A		
Are there any sub-studies linked to this study?	☐ Yes ☑ No		
If yes, could you please provide the HREC Ref's for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	Professor Landon Myer		
Department / Office Internal Mail Address	Level 5, School of Public Health and Family Medicine, Falmouth Building, Faculty of Health Sciences		

1.1 Does this protocol receive US Federal funding?	☑ Yes	☐ No

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

9.1.2 University of Cape Town Human Research Ethics Committee for the B-Positive study (all components)

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
Email: nosi.tsame@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

13 November 2015

HREC REF: 749/2015

A/Prof A Boulle
CIDER
Public Health & Family Medicine
Falmouth Building

Dear A/Prof Boulle

PROJECT TITLE: A POPULATION BASED EVALUATION OF EXPANDED (ANTI-RETROVIRAL THERAPY) ART ACCESS IN PREGNANCY ('B POSITIVE') Sub-study linked to 541/2015

Thank you for your letter to the Faculty of Health Sciences Human Research Ethics Committee dated 06 November 2015.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30*November 2016.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Yours sincerely

Signed by candidate

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI),and Declaration of Helsinki guidelines.

HREC 749/2015

13.11.15

HUMAN RESEARCH ETHICS COMMITTEE
17 AUG 2021
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938)			
This serves as notification of annual approval, including any documentation described below.			
☐ Approved	Annual progress report	Approved until/next renewal date	30·8·22
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee	Signed by candidate	Date Signed	23/8/2021

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	21 July 2021		
HREC REF Number	749/2015	Current Ethics Approval was granted until	30/08/2021
Protocol title	B positive: a population-based evaluation of expanded (anti-retroviral therapy) ART access in pregnancy		
Protocol number (if applicable)	5.0		
Are there any sub-studies linked to this study?	x Yes	☐ No	
If yes, could you please provide the HREC Reference number for all sub-studies? **Note**: A separate FHS016 must be submitted for each sub-study.	HREC REF: 541/2015 (separate FHS016 enclosed)		
Principal Investigator	Professor Andrew Boulle changing to Dr Emma Kalk (FHS002 enclosed)		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

9.1.3 University of Cape Town Human Research Ethics Committee for this thesis

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room G50- Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6492
Email: hrec-enquiries@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

24 February 2020

HREC REF: 019/2020

Prof L Myer
School of Public Health & Family Medicine
Level -5
Falmouth Building-FHS

Dear Prof Myer

PROJECT TITLE: METHODOLOGICAL CONCERNS RELATED TO THE EPIDEMIOLOGIC STUDY BIRTH OUTCOMES (SUB-STUDY - 739/2014) (MS THOKOZILE ROSEMARY MALABA)

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee (HREC) for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 28 February 2021

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

The HREC acknowledge that the student: - Ms Thokozile Malaba will also be involved in this study.

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

Yours sincerely

Signed by candidate

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE

Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

HREC 019/2020sa

HUMAN RESEARCH ETHICS COMMITTEE
24 MAY 2022
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

UNIVERSITY OF CAPE TOWN
FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☐ Approved	Annual progress report	Approved until/next renewal date	30.5.23
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee	Signed by candidate	Date Signed	24/5/2022

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website: http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	20th May 2022		
HREC REF Number	019/2020	Current Ethics Approval was granted until	28/02/2022
Protocol title	Methodological concerns related to the epidemiologic study of birth outcomes		
Protocol number (if applicable)	N/A		
Are there any sub-studies linked to this study?	☐ Yes	☑ No	
If yes, could you please provide the HREC Reference number for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.			
Principal Investigator	Professor Landon Myer		
Department / Office Internal Mail Address	Level 5, School of Public Health and Family Medicine, Falmouth Building, Faculty of Health Sciences.		

1.1 Does this protocol receive US Federal funding?	☐ Yes	☑ No

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

9.2 Supplementary materials for Chapter 4

Supplementary Table 9.2.1a Description of women with all three GA assessment methods (n=629)

	Total n=629	HIV-uninfected n=309	HIV-infected n=320	HIV infected	
				Initiated before pregnancy n=134	Initiated during pregnancy n=186
Maternal Characteristics					
Age, years					
<24	173 (28)	109 (35)	64 (20)	18 (13)	46 (25)
25-29	164 (26)	94 (30)	70 (22)	18 (13)	52 (28)
>30	292 (46)	106 (34)	186 (58)	98 (74)	88 (47)
Median	29 (24-33)	27 (23-31)	31 (26-34)	33 (29-36)	29 (25-32)
Height, cm					
≤155	188 (30)	92 (30)	96 (30)	34 (25)	62 (33)
156-161	232 (37)	111 (36)	121 (38)	57 (43)	64 (35)
≥162	209 (33)	106 (34)	103 (32)	43 (32)	60 (32)
Median	158 (154-163)	158 (154-163)	158 (154-162)	158 (155-163)	157 (153-162)
Gestation at enrolment, weeks					
Determined by LMP					
Median	17 (14-21)	18 (15-21)	17 (13-20)	16 (12-20)	17 (13-20)
Determined by SFH					
Median	20 (16-23)	21 (17-24)	18 (15-22)	18 (14-21)	19 (15-23)
Determined by US					
Median	19 (15-23)	20 (17-24)	18 (14-22)	18 (14-21)	19 (14-22)
Body Mass Index					
Underweight	6 (1)	5 (2)	1 (0.3)	0	1 (0.5)
Normal	122 (19)	57 (18)	65 (20)	25 (19)	40 (24)
Overweight	178 (28)	86 (28)	92 (29)	39 (31)	53 (28)
Obese	290 (46)	146 (47)	144 (45)	63 (50)	81 (44)
Gravidity					
1	127 (20)	78 (25)	49 (15)	16 (12)	33 (18)
2	231 (37)	115 (37)	116 (36)	43 (33)	72 (39)
≥3	271 (43)	116 (38)	155 (48)	70 (54)	81 (44)
Median	2 (2-3)	2 (1-3)	2 (2-3)	3 (2-3)	2 (2-3)
Parity					
0	153 (24)	88 (28)	65 (20)	23 (17)	42 (23)
1	247 (39)	121 (39)	126 (39)	49 (37)	77 (41)
≥2	229 (37)	100 (33)	129 (40)	62 (46)	67 (36)
Median	1 (1-2)	1 (0-2)	1 (1-2)	1 (1-2)	1 (1-2)
Previous Preterm*					
Yes	44 (9)	18 (8)	26 (10)	12 (10)	14 (9)
Haemoglobin g/dl					
Normal (≥11.0)	260 (41)	150 (49)	110 (34)	56 (43)	51 (28)
Mild Anaemia (9-10.9)	233 (37)	104 (34)	129 (40)	44 (34)	82 (46)
Moderate Anaemia (7-8.9)	39 (6)	11 (4)	28 (9)	10 (8)	16 (9)
Severe Anaemia (<7g)	1 (0.2)	1 (0.3)	0	0	0
ART Regimen					
TDF-3TC-EFV	290 (46)	-	290 (95)	111 (93)	179 (97)
TDF-3TC-NVP	2 (0.3)	-	2 (1)	0	2 (1)
Other NNRTI	6 (1)	-	6 (2)	4 (3)	2 (1)
PI regimen	7 (1)	-	7 (2)	5 (4)	2 (1)

*among women with a previous pregnancy

All variables, with the exception of height, BMI and haemoglobin had <1% missing data. For height, 5% (n=30) of data was missing and for BMI, 5% (n=34) was missing, for both there were similar proportions across all comparison groups. For Haemoglobin, 15% (n=96) of data was missing with similar proportions across all comparison groups. For ART status 2% (n=15) were missing regimen data

Supplementary Table 9.2.1b Description of women with all three GA assessment methods (n=629) vs women without all three GA assessment methods (n=1158)

	Overall Cohort n=1787	<3 measures n=1158	3 measures n=629
Maternal Characteristics			
Age, years			
<24	536 (30)	363 (31)	173 (28)
25-29	500 (28)	336 (29)	164 (26)
>30	751 (42)	459 39)	292 (46)
Median	28 (24-32)	28 (23-32)	29 (24-33)
Height, cm			
≤155	500 (28)	312 (27)	188 (30)
156-161	651 (36)	419 (36)	232 (37)
≥162	636 (36)	427 (37)	103 (32)
Median	158 (155-163)	159 (155-163)	158 (154-163)
Gestation at enrolment, weeks			
Determined by LMP	1585	956	629
17 (12-22)	17 (14-21)	16 (11-25)	17 (14-21)
Determined by SFH	1219	590	629
23 (18-28)	20 (16-23)	27 (22-231)	20 (16-23)
Determined by US	1204	575	629
16 (12-21)	19 (15-23)	13 (9-17)	19 (15-23)
Body Mass Index			
Underweight	16 (1)	10 (1)	6 (1)
Normal	329 (18)	207 (19)	122 (20)
Overweight	494 (28)	317 (30)	178 (30)
Obese	823 (46)	533 (50)	290 (49)
Gravidity			
1	383 (21)	256 (22)	127 (20)
2	639 (36)	408 (35)	231 (37)
≥3	765 (43)	494 (43)	271 (43)
Median	2 (2-3)	2 (2-3)	2 (2-3)
Parity			
0	472 (26)	319 (28)	153 (24)
1	700 (39)	453 (39)	247 (39)
≥2	615 (35)	386 (33)	229(36)
Median	1 (0-2)	1 (0-2)	1 (1-2)
Previous Preterm*			
Yes	127 (9)	83 (7)	44 (7)
Haemoglobin g/dl			
Normal (≥11.0)	792 (44)	532 (53)	260 (49)
Mild Anaemia (9-10.9)	613 (34)	380 (38)	233 (44)
Moderate Anaemia (7-8.9)	122 (7)	83 (8)	39 (7)

	Overall Cohort n=1787	<3 measures n=1158	3 measures n=629
Severe Anaemia (<7g)	3 (0.2)	2 (0.2)	1 (0.2)
HIV Status			
HIV-uninfected	1014 (57)	705 (61)	309 (49)
HIV-infected	773 (43)	453 (39)	320 (51)
ART Status			
Initiated before pregnancy	368 (48)	234 (52)	134 (42)
Initiated during pregnancy	405 (52)	219 (19)	186 (58)
ART Regimen			
TDF-3TC-EFV	691 (39)	401 (35)	290 (46)
TDF-3TC-NVP	4 (0.2)	2 (0.2)	2 (0.3)
Other NNRTI	20 (1)	14 (1)	6 (1)
PI regimen	22 (1)	15 (1)	7 (1)

*among women with a previous pregnancy
All variables, with the exception of height, BMI and haemoglobin had <1% missing data. For height, 6% (n=111) of data was missing and for BMI, 7% (n=124) was missing, for both there were similar proportions across all comparison groups. For Haemoglobin, 14% (n=257) of data was missing with similar proportions across all comparison groups. For ART status 2% (n=36) were missing regimen data

Supplementary Table 9.2.2 Birth outcomes in women with all three GA assessment methods (n=629)

	Total n=629	HIV-uninfected n=309	HIV-infected n=320	HIV-infected	
				Initiated before pregnancy n=134	Initiated during pregnancy n=186
Gestational Age (weeks) LMP					
Post-term >42	36 (6)	15 (5)	21 (7)	12 (9)	9 (5)
Term (37-42)	329 (52)	159 (51)	170 (53)	70 (52)	100 (54)
Any Preterm (<37)*	264 (42)	135 (44)	129 (40)	52 (39)	77 (41)
Late to Mod Preterm (32-37)	*186 (70)*	*101 (75)*	*85 (66)*	*32 (62)*	*53 (69)*
Very Preterm (28-32)	*78 (30)*	*34 (25)*	*44 (34)*	*20 (38)*	*24 (31)*
Gestational Age (weeks) SFH					
Post-term >42	42 (7)	24 (8)	18 (6)	6 (5)	12 (7)
Term (37-42)	498 (79)	247 (80)	251 (78)	104 (83)	147 (85)
Any Preterm (<37)*	89 (14)	38 (12)	51 16)	24 (17)	27 (15)
Late to Mod Preterm (32-37)	*68 (76)*	*32 (84)*	*36 (71)*	*20 (83)*	*16 (59)*
Very Preterm (28-32)	*21 (24)*	*6 (16)*	*15 (29)*	*4 (17)*	*11 (41)*
Gestational Age (weeks) US					
Post-term >42	3 (0.5)	3 (1)	0	0	0
Term (37-42)	565 (90)	285 (92)	280 (88)	121 (90)	159 (85)
Any Preterm (<37)*	61 (10)	21 (7)	40 (12)	13 (10)	27 (15)
Late to Mod Preterm (32-37)	*48 (79)*	*19 (90)*	*29 (73)*	*9 (69)*	*20 (74)*
Very Preterm (28-32)	*13 (21)*	*2 (10)*	*11 (27)*	*4 (31)*	*7 (26)*
Size for Gestational Age LMP	n=578	n=287	n=291	n=117	n=174
Small (SGA)	51 (9)	17 (6)	34 (12)	12 (10)	22 (13)
Appropriate (AGA)	320 (55)	159 (55)	161 (55)	68 (58)	93 (53)
Large (LGA)	208 (36)	111 (39)	96 (33)	37 (32)	59 (34)
Size for Gestational Age SFH	n=581	n=284	n=297	n=125	n=172
Small (SGA)	84 (14)	33 (11)	51 (17)	14 (12)	37 (22)
Appropriate (AGA)	440 (76)	224 (79)	216 (73)	99 (80)	117 (68)
Large (LGA)	57 (10)	27 (10)	30 (10)	12 (8)	18 (10)
Size for Gestational Age US	n=618	n=305	n=313	n=130	n=183
Small (SGA)	65 (11)	22 (7)	43 (14)	17(13)	26 (14)
Appropriate (AGA)	505 (82)	258 (85)	247 (79)	105 (81)	142 (78)
Large (LGA)	48 (7)	25 (8)	23 (7)	8 (6)	15 (8)

* Any Preterm consists of the subset of Late Preterm, Moderately Preterm and Very Preterm deliveries
For Size for GA data missing for women whose GA is not within Intergrowth standard limits (<24w and >42w) and those missing either GA or birthweight

Supplementary Table 9.2.3 Adjusted association between HIV/ART status and preterm delivery in women with singleton deliveries (n=1787)

Assessment Method	**HIV-infected vs HIV-uninfected (Ref)***			**Initiated before pregnancy (Ref) vs Initiated during pregnancy****		
		AOR (95% CI)	**P-value**		**AOR (95% CI)**	**P-value**
Last menstrual period (LMP) (n=1470)	HIV-infected	1.12 (0.90-1.39)	0.325	During pregnancy	0.81 (0.57-1.13)	0.252
Symphysis Fundal Height (SFH) (n=1122)	HIV-infected	1.07 (0.78-1.46)	0.666	During pregnancy	1.15 (0.70-1.89)	0.587
Ultrasound (US) (n=1198)	HIV-infected	1.53 (1.04-2.26)	0.030	During pregnancy	0.79 (0.47-1.33)	0.371

*Adjusted for age, parity, BMI and previous PTD

**Adjusted for age, parity, BMI, previous PTD and ART regimen

Supplementary Table 9.2.4a Gestational Age at 1st ANC visit and delivery according to different assessment methods women with all three GA assessment methods (n=629): HIV and ART status

Assessment Method	**Total n=629**	**HIV-uninfected n=309**	**HIV-infected n=320**	**HIV-infected n= 320**	
				Initiated before pregnancy n=134	**Initiated during pregnancy n=186**
LMP					
Median GA*	17 (14-21)	18 (15-21)	17 (13-20)	16 (12-20)	17 (13-20)
Median GA**	38 (34-39)	37 (34-39)	38 (34-39)	38 (34-40)	37 (34-39)
SFH					
Median GA*	20 (16-23)	21 (17-24)	18 (15-22)	18 (14-21)	19 (15-23)
Median GA**	39 (38-40)	39 (38-40)	39 (37-40)	39 (37-40)	39 (38-40)
US					
Median GA*	19 (15-23)	20 (17-24)	18 (14-22)	18 (14-21)	19 (14-22)
Median GA**	39 (38-40)	39 (38-40)	39 (38-40)	39 (38-40)	39 (38-40)

*Gestation at 1st ANC Visit
**Gestation at delivery

Supplementary Table 9.2.4b Gestational Age at 1st ANC visit and delivery according to different assessment methods women with all three GA assessment methods (n=629): BMI status

Assessment Method	Underweight n=6	Normal n=122	Overweight n=178	Obese n=290	Unknown n=33
LMP					
Median GA*	19 (17-20)	17 (14-19)	18 (15-21)	17 (14-21)	17 (13-20)
Median GA**	36 (33-38)	36 (33-39)	38 (34-40)	38 (34-39)	37 (35-39)
SFH					
Median GA*	22 (19-23)	19 (15-23)	20 (16-23)	20 (15-23)	21 (15-22)
Median GA**	37 (36-38)	39 (37-40)	39 (38-40)	39 (38-40)	39 (38-40)
US					
Median GA*	23 (21-25)	19 (16-22)	19 (16-23)	20 (15-23)	19 (14-22)
Median GA**	38 (37-39)	38 (37-39)	39 (38-40)	39 (38-40)	38 (38-39)

*Gestation at 1st ANC Visit
**Gestation at delivery

Supplementary Table 9.2.5 Adjusted association between HIV/ART status and preterm delivery in women with all three assessment methods (n=629)

Assessment Method	**HIV-infected vs HIV-uninfected (Ref)**			**Initiated before pregnancy (Ref) vs Initiated during pregnancy**		
		aOR (95% CI)	**P-value**		**aOR (95% CI)**	**P-value**
Last Menstrual Period (LMP)	HIV-infected	0.90 (0.64-1.26)	0.532	During pregnancy	0.92 (0.55-1.55)	0.757
Symphysis Fundal Height (SFH)	HIV-infected	1.34 (0.84-2.15)	0.216	During pregnancy	1.37 (0.72-2.63)	0.341
Ultrasound (US)	HIV-infected	2.00 (1.13-3.53)	0.017	During pregnancy	0.67 (0.32-1.43)	0.303

*Adjusted for age, parity, BMI

**Adjusted for age, parity, BMI, and ART regimen

9.3 Supplementary materials for Chapter 5

Supplementary Table 9.3.1 Preterm delivery misclassification bias

PART A: Ultrasound based preterm delivery status contingency table*

	HIV+	**HIV-**
PTD+	A	B
PTD-	C	D

* ultrasound considered gold standard therefore ultrasound-based measure is assumed to give the true classification
PTD – preterm delivery

PART B: LMP/SFH based preterm delivery status contingency table*

	HIV+	**HIV-**
PTD+	a	b
PTD-	c	d
Total	*a+c* HIV^{+}_{Total}	*b+d* HIV^{-}_{Total}
Sensitivity (SE) of assessment method for true outcome status = a/(a+c)		
Specificity (SP) of assessment method for true outcome status = b(b+d)		

* Observed preterm delivery status with biased classification scheme known, with bias parameters (SE and SP) calculated

Part C: Corrected contingency table cell values

	HIV+	**HIV-**
PTD+	a (corrected) $\frac{[a - HIV^{+}_{Total}(1 - SP_{HIV+})]}{[SE_{HIV+} - (1 - SP_{HIV+})]}$	b (corrected) $\frac{[b - HIV^{-}_{Total}(1 - SP_{HIV-})]}{[SE_{HIV-} - (1 - SP_{HIV-})]}$
PTD-	c (corrected) HIV^{+}_{Total} – c (corrected)	d (corrected) HIV^{-}_{Total} – d (corrected)
Corrected Risk Ratio: $\frac{a\ (corrected)/HIV^{+}_{Total}}{b\ (corrected)/HIV^{-}_{Total}}$		

Supplementary Table 9.3.2 Imputed baseline characteristics of all enrolled pregnant women at 1st antenatal care visit in variables with missing data >5% (n=3952)

	WLHIV N = 1445	**HIV-negative N = 2507**	**WLHIV N= 1445**	
			Initiated ART before pregnancy N = 717	**Initiated ART during pregnancy N = 728**
Height (cm)				
≤155	433 (30)	707 (28)	208 (29)	225 (31)
156-161	609 (42)	1104 (44)	308 (43)	301 (41)
≥162	403 (28)	695 (28)	201 (28)	202 (28)
Median (IQR)	158 (154, 162)	158 (154, 162)	158 (154, 162)	158 (154, 162)
Haemoglobin (g/dl)				
Normal (≥11.0)	713 (50)	1430 (57)	408 (57)	305 (42)
Mild anaemia (9.0-10.9)	556 (38)	876 (35)	234 (33)	322 (44)
Moderate anaemia (7.0-8.9)	175 (12)	197 (8)	75 (10)	100 (14)
Severe anaemia (<7.0)	1 (0.1)	5 (0.2)	0	1 (0.1)
Median (IQR)	11 (10, 12)	11 (10, 12)	11 (10, 12)	11 (10, 12)
Body Mass Index (kg/m^2)				
Normal (≤24.9)	305 (21)	532 (21)	141 (20)	162 (23)
Overweight (25.0-29.9)	420 (29)	684 (27)	217 (30)	203 (27)
Obese (≥30.0)	720 (50)	1291 (52)	358 (50)	363 (50)
Median (IQR)	30 (26, 35)	30 (26, 36)	30 (26, 35)	30 (25, 35)

WLHIV – women living with HIV, ART – antiretroviral therapy

Supplementary Table 9.3.3 Gestational age at 1st antenatal care visit according to different assessment methods women by HIV status and timing of ART initiation

Assessment method	**WLHIV**	**HIV-negative**	**WLHIV**	
			Initiated ART before pregnancy	**Initiated ART during pregnancy**
Last menstrual period	1259	2216	643	616
Median gestational age (weeks)	16 (11, 22)	17 (12, 23)	15 (11, 22)	17 (12, 22)
Symphysis fundal height	924	1509	414	510
Median gestational age (weeks)	22 (17, 27)	23 (19, 28)	22 (16, 27)	23 (18, 28)
Ultrasound	929	1411	437	492
Median gestational age (weeks)	15 (11, 20)	16 (12, 21)	14 (10, 19)	17 (12, 21)

WLHIV – women living with HIV, ART – antiretroviral therapy

Supplementary Table 9.3.4 Baseline characteristics of all enrolled pregnant women at 1st antenatal care visit by assessment method (n=3885)

	Ultrasound N = 2340	**Symphysis fundal height N = 1151**	**Last menstrual period N = 394**
Age (years)			
≤24	690 (30)	483 (38)	129 (33)
25-29	678 (29)	295 (26)	111 (28)
≥30	972 (42)	418 (36)	154 (39)
Median (IQR)	28 (24, 32)	27 (22, 32)	27 (23, 33)
Height (cm)			
≤155	717 (31)	349 (30)	104 (26)
156-161	834 (36)	394 (34)	125 (32)
≥162	621 (27)	315 (27)	123 (31)
Missing	*168 (7)*	*93 (8)*	*42 (11)*
Median (IQR)	158 (154, 162)	158 (154, 162)	159 (155, 163)
Body Mass Index (kg/m^2)			
Normal (≤24.9)	468 (20)	195 (17)	83 (21)
Overweight (25.0-29.9)	579 (25)	339 (29)	88 (22)
Obese (>30.0)	1108 (47)	519 (45)	178 (45)
Missing	*185 (8)*	*98 (9)*	*45 (11)*
Median (IQR)	30 (26, 35)	30 (26, 35)	30 (25, 36)
Gravidity			
1	551 (24)	296 (26)	109 (28)
2	822 (35)	387 (34)	128 (33)
≥3	930 (40)	467 (41)	156 (40)
Missing	*37 (2)*	*1 (0.1)*	*1 (0.3)*
Parity			
0	714 (31)	331 (29)	136 (35)
1	884 (38)	428 (37)	138 (35)
≥2	704 (31)	391 (34)	119 (30)
Missing	*38 (2)*	*1 (0.1)*	*1 (0.3)*
Previous preterm*			
Yes	187 (11)	74 (9)	24 (9)
Haemoglobin (g/dl)			
Normal (≥11.0)	1096 (47)	393 (34)	195 (50)
Mild anaemia (9.0-10.9)	680 (29)	404 (35)	72 (18)
Moderate anaemia (7.0-8.9)	132 (6)	105 (9)	12 (3)
Severe anaemia (<7.0)	2 (0.1)	2 (0.2)	1 (0.3)
Missing	*430 (18)*	*247 (22)*	*114 (29)*
Gestational age at 1st antenatal care visit* (weeks)			
Median (IQR)	16 (11, 21)	26 (22, 31)	13 (9, 19)
Booking blood pressure (mmHG) (%)			
Normal (Systolic <120 or Diastolic <80)	1533 (63)	714 (63)	202 (63)
Prehypertensive (Systolic 120-139 and Diastolic 80-89)	672 (31)	383 (31)	255 (31)
Hypertensive (Systolic >140 or Diastolic >90)	85 (4)	47 (4)	28 (4)
Missing	*50 (2)*	*7(0.6)*	*9 (2)*
Any hypertension**	242 (10)	103 (9)	34 (9)
HIV Status			
Negative	1411 (60)	779 (68)	279 (71)
Positive	929 (40)	372 (32)	115 (29)
ART initiated before pregnancy	437 (19)	190 (16)	79 (20)
ART initiated during pregnancy	492 (21)	182 (16)	36 (9)

ART – antiretroviral therapy

* best available measure: ultrasound measure used when available, followed by symphysis fundal height then bias corrected last menstrual period.

** hypertension at any point in pregnancy

Supplementary Table 9.3.5 Gestational length according to different assessment methods women by HIV status and timing of ART initiation

	WLHIV	HIV-negative	WLHIV	
			Initiated ART before pregnancy	Initiated ART during pregnancy
Gestational age (weeks) LMP	1028	1857	526	502
Post-term >42	46 (5)	102 (5)	22 (4)	24 (5)
Term (37-42)	558 (54)	1130 (61)	301 (57)	257 (51)
Any preterm (<37)*	424 (41)	625 (34)	203 (39)	221 (44)
Gestational age (weeks) SFH	811	1338	357	454
Post-term >42	53 (6)	121 (9)	22 (6)	31 (7)
Term (37-42)	591 (73)	247 (72)	266 (75)	325 (72)
Any preterm (<37)*	167 (21)	252 (19)	69 (19)	98 (22)
Gestational age (weeks) US	813	1210	382	431
Post-term >42	0	8 (0.7)	0	0
Term (37-42)	690 (85)	1031 (85)	325 (85)	365 (85)
Any preterm (<37)*	123 (15)	171 (14)	57 (15)	66 (15)

WLHIV – women living with HIV, ART – antiretroviral therapy
LMP – last menstrual period, SFH – symphysis fundal height, US - ultrasound
* Any preterm consists of the subset of late preterm, moderately preterm and very preterm deliveries

Supplementary Table 9.3.6 Misclassification bias correction by quantitative bias adjustment

Assessment Method		**Conventional RR (95% CI)**	**QBA (overall bias parameters)**	**QBA (strata specific parameters) ND***	**QBA (strata specific parameters) D****
Last menstrual period (n=2716)	HIV-negative	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	Living with HIV	1.21 (1.10, 1.33)	1.35 (1.22, 1.5)	1.44 (1.28, 1.62)	1.35 (1.11, 1.64)
			Sensitivity = 63.38	HIV+ Sensitivity = 60.61	
				HIV- Sensitivity = 67.50	
			Specificity = 86.32	HIV+ Specificity = 82.32	
				HIV- Specificity = 89.24	
Symphysis fundal height (n=1961)	HIV-negative	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
	Living with HIV	1.06 (0.89, 1.27)	1.09 (0.91, 1.30)	1.08 (0.93, 1.28)	0.79 (0.59, 1.07)
			Sensitivity = 76.92	HIV+ Sensitivity = 87.50	
				HIV- Sensitivity = 60.00	
			Specificity = 95.06	HIV+ Specificity = 95.65	
				HIV- Specificity = 94.29	
Ultrasound (n=1981)	HIV-negative	1.00 (reference)			
	Living with HIV	1.06 (0.86, 1.32)			

QBA – quantitative bias analysis, RR – risk ratio
*ND – nondifferential
**D – differential

Supplementary Table 9.3.7 Birth outcomes in women living with HIV by timing of ART initiation (n=1424)

	WLHIV N = 1424			
	Initiated ART before pregnancy N = 710	**Initiated ART during pregnancy 1st trimester N = 197**	**Initiated ART during pregnancy 2nd trimester N = 377**	**Initiated ART during pregnancy 3rd trimester N = 126**
Pregnancy outcome				
Live births	622 (91)	157 (82)	344 (95)	118 (98)
Miscarriage	30 (4)	17 (9)	8 (2)	0
Stillbirth	13 (2)	6 (3)	9 (3)	3 (2)
Nonviable	16 (2)	6 (3)	2 (0.6)	0
ξ Other losses	2 (0.3)	5 (3)	0	0
Missing	*27 (4)*	*26 (4)*	*26 (4)*	*26 (4)*
	622	157	344	118
Gestational length (weeks)				
Post-term (≥42)	23 (4)	4 (3)	5 (2)	5 (4)
Term (37-41)	491(79)	125 (80)	276 (80)	87 (74)
Preterm (<37)*	108 (17)	28 (19)	63 (18)	26 (22)
Late/moderate preterm (32-<37)	84 (14)	22 (14)	52 (15)	24 (16)
Very preterm (28-31)	19 (3)	6 (4)	5 (1)	2 (2)
Extremely preterm (<28)	3 (0.5)	0	6 (1)	0
Clinical presentation at labour				
Spontaneous	54 (9)	13 (10)	35 (10)	16 (10)
Clinician-initiated	48 (7)	14 (8)	25 (8)	9 (8)
Unknown	6 (1)	1 (1)	3 (1)	1 (1)
Missing	*6 (1)*	*2 (2)*	*2 (1)*	*0*
Size for gestational age (centile)				
LGA (>90th)	82 (13)	31 (20)	47 (14)	15 (13)
AGA (10th – 90th)	442 (71)	104 (66)	247 (72)	90 (76)
SGA (<10th)	72 (12)	16 (10)	41 (12)	8 (7)
Missing	*26 (4)*	*6 (4)*	*9 (2)*	*5 (4)*

WLHIV – women living with HIV, ART – antiretroviral therapy
LGA – Large for Gestational Age, AGA – Appropriate for Gestational Age, SGA – Small for Gestational Age
* preterm deliveries not classified because there was no defined gestational age (n=2)

Supplementary Table 9.3.8 Pregnancy outcomes by HIV status and timing of ART initiation status among enrolled women with singleton pregnancies with imputed outcome data (n=3894)

	WLHIV N = 1424	**HIV-negative N = 2470**	**WLHIV N = 1424**	
			Initiated ART before pregnancy N = 710	**Initiated ART during pregnancy N = 714**
Pregnancy outcome				
Miscarriage	67 (5)	114 (5)	37 (5)	31 (4)
Stillbirth	37 (3)	74 (3)	15 (2)	22 (3)
Nonviable	34 (2)	15 (0.6)	19 (3)	14 (2)
Other[ξ]	10 (0.7)	12 (0.5)	3 (0.4)	7 (1)
Live Births	1276 (90)	2255 (91)	636 (90)	640 (90)

WLHIV – women living with HIV, ART – antiretroviral therapy

ξ other losses include abortion, ectopic pregnancy and maternal death

Supplementary Table 9.3.9 Birthweight outcomes by HIV status and timing of ART initiation among enrolled women with singleton pregnancies (n=3426)

	WLHIV N = 1247	HIV-negative N = 2179	WLHIV N = 1247			
			Initiated ART before pregnancy N = 625	Initiation during pregnancy N = 622		
				Initiated ART during pregnancy 1st trimester N = 197	Initiated ART during pregnancy 2nd trimester N = 377	Initiated ART during pregnancy 3rd trimester N = 126
Birthweight (grams)						
High (≥4000)	45 (4)	108 (5)	16 (3)	4 (3)	23 (7)	2 (2)
Normal (2500-3999)	1011 (81)	1795 (82)	513 (83)	127 (80)	260 (76)	107 (91)
Low (<2500)	175 (14)	253 (12)	91 (14)	18 (12)	54 (16)	8 (7)
LBW (2500-1500)	153 (12)	216 (10)	79 (12)	16 (11)	38 (11)	8 (7)
Very LBW (<1500)	22 (2)	37 (2)	12 (2)	2 (1)	8 (2)	0
Mean (SD)	3061 (590)	3154 (594)	3042 (578)	3097 (592)	3050 (640)	3157 (466)
Missing	*16 (1)*	*23 (1)*	*2 (0.3)*	*6 (4)*	*7 (2)*	*1 (1)*

WLHIV – women living with HIV, ART – antiretroviral therapy
LBW – low birthweight

Supplementary Table 9.3.10 Associations between HIV and ART exposure groups and low birthweight among women with live singleton deliveries (n=3426)

Outcome measure	Comparison A†			Comparison B††			Comparison C†††			Comparison D††††		
		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)
Low birth weight (<2500g)*	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.21 (0.99, 1.47)	1.13 (0.91, 1.39)	During pregnancy	0.94 (0.70, 1.26)	0.91 (0.66, 1.25)	Before pregnancy	1.25 (0.98, 1.59)	1.19 (0.92, 1.54)	Before pregnancy	1.25 (0.98, 1.59)	1.19 (0.92, 1.54)
							During pregnancy	1.17 (0.92, 1.50)	1.07 (0.82, 1.39)	1st trimester	1.13 (0.72, 1.78)	0.93 (0.57, 1.53)
										2nd trimester	1.35 (1.02, 1.83)	1.34 (0.99, 1.83)
										3rd trimester	0.58 (0.99, 1.83)	0.46 (0.22, 0.98)

WLHIV – women living with HIV, ART – antiretroviral therapy, RR – risk ratio, ARR – adjusted risk ratio, CI – confidence interval

†Comparison A (WLHIV vs HIV-negative)

††Comparison B (WLHIV ART initiated during pregnancy vs WLHIV ART initiated before conception)

†††Comparison C (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy vs HIV-negative)

††††Comparison D (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy (1st trimester) vs WLHIV ART initiated during pregnancy (2nd trimester) vs WLHIV ART initiated during pregnancy (3rd trimester) vs HIV-negative)

* adjusted for age, maternal height, hypertension, BMI, gestational age at delivery

Supplementary Table 9.3.11 Associations between HIV and ART exposure groups and preterm delivery among women with live singleton deliveries by assessment method

Preterm Delivery (<37 weeks)	Comparison A†			Comparison B††			Comparison C†††			Comparison D††††		
		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)		RR (95% CI)	ARR (95% CI)
Last menstrual period*	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.21 (1.07, 1.37)	1.19 (1.07, 1.31)	During pregnancy	1.15 (0.94, 1.39)	1.14 (0.98, 1.33)	Before pregnancy	1.13 (0.97, 1.33)	1.11 (0.97, 1.26)	Before pregnancy	1.13 (0.97, 1.33)	1.10 (0.97, 1.26)
							During pregnancy	1.30 (1.11, 1.51)	1.27 (1.12, 1.42)	1st trimester	0.80 (0.57, 1.12)	0.81 (0.61, 1.08)
										2nd trimester	1.41 (1.17, 1.70)	1.36 (1.18, 1.56)
										3rd trimester	1.66 (1.25, 2.20)	1.60 (1.33, 1.94)
Symphysis fundal height*	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.06 (0.87, 1.29)	1.08 (0.89, 1.32)	During pregnancy	1.12 (0.83, 1.53)	1.08 (0.82, 1.43)	Before pregnancy	0.99 (0.76, 1.30)	1.02 (0.80, 1.30)	Before pregnancy	0.99 (0.76, 1.30)	1.02 (0.80, 1.30)
										1st trimester	0.79 (0.37, 1.67)	0.80 (0.40, 1.59)
							During pregnancy	1.12 (0.89, 1.41)	1.13 (0.92, 1.39)	2nd trimester	1.15 (0.88, 1.51)	1.17 (0.92, 1.49)
										3rd trimester	1.17 (0.78, 1.76)	1.17 (0.83, 1.65)
Ultrasound*	HIV-negative	1.00 (reference)	1.00 (reference)	Before pregnancy	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)	HIV-negative	1.00 (reference)	1.00 (reference)
	WLHIV	1.06 (0.84, 1.34)	0.98 (0.78, 1.23)	During pregnancy	1.03 (0.72, 1.46)	1.05 (0.75, 1.49)	Before pregnancy	1.05 (0.78, 1.42)	0.94 (0.70, 1.26)	Before pregnancy	1.05 (0.78, 1.42)	0.94 (0.70, 1.26)
							During pregnancy	1.08 (0.81, 1.43)	1.02 (0.70, 1.33)	1st trimester	1.05 (0.67, 1.65)	0.96 (0.62, 1.47)
										2nd trimester	1.14 (0.81, 1.59)	1.09 (0.79, 1.49)
										3rd trimester	0.70 (0.22, 2.20)	0.72 (0.24, 2.20)

WLHIV – women living with HIV, ART – antiretroviral therapy, RR – risk ratio, ARR – adjusted risk ratio, CI – confidence interval

†Comparison A (WLHIV vs HIV-negative) ††Comparison B (WLHIV ART initiated during pregnancy vs WLHIV ART initiated before conception) †††Comparison C (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy vs HIV-negative)

††††Comparison D (WLHIV ART initiated during pregnancy vs WLHIV ART initiated during pregnancy (1st trimester) vs WLHIV ART initiated during pregnancy (2nd trimester) vs WLHIV ART initiated during pregnancy (3rd trimester) vs HIV-negative)

* adjusted for age, parity, hypertension

Supplemental File 9.3.1 Quantitative Bias Analyses

To explore the impact of a range of misclassification probabilities on our association of interest we undertook probabilistic bias analysis (PBA), which enables values assigned to the bias parameter to be varied (1). We chose the triangular distribution as the closest to a normal distribution, and because there was no evidence to suggest that the data was not normally distributed. This distribution was assigned to each bias parameter using calculated sensitivity and specificity as the mode and the upper and lower bounds as ±0.50 of the mode. Each parameter was randomly sampled 1000 times, calculating 1000 risk ratios (RR), summarised as the median RR and 95% simulation interval (SI), the 2.5th and 97.5th percentile of the RR distribution.

To establish if LMP- and SFH-based outcome misclassification was non-differential we conducted three QBA analyses. The initial QBA iteration ('overall bias parameters') used one overall sensitivity and specificity for the entire LMP/SFH cohort; subsequent analyses used 'strata-specific bias parameters' with separate sensitivity and specificity parameters by HIV status with misclassification assumed to be non-differential or differential (2).

References

1. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data: Springer Science & Business Media; 2011.
2. Walraven C. A comparison of methods to correct for misclassification bias from administrative database diagnostic codes. *International Journal of Epidemiology.* 2018;47(2):605-16.

Supplementary File 9.3.2 Multiple overimputation

Multiple overimputation is an extension of multiple imputation (MI) of missing values (in partially observed variables) to the mitigation of measurement error. Multiple overimputation, an alternative to estimation-based error-correction approaches (1), allows data values to be placed on a continuum ranging from observed without error (perfect measurement) to increasing levels of measurement error and missingness to no measurement at all (2). This method treats measurement error as partially missing information and completely missing values as an extreme form of measurement error. Missing values are multiply imputed based on observed covariates. Mismeasured values are used to generate observation-level Bayesian priors for the imputation model and are then replaced (overimputed) with multiply imputed values. The priors are specified to incorporate prior knowledge (in the form of the observed mismeasured value), as well as appropriate uncertainty into the imputation model and can be estimated within the procedure from gold standard or proxy variables (3).

In our study, the overimputation model contained variables to be used in the analysis model which included the exposure (HIV status and timing of ART initiation), gestation-based outcomes (preterm delivery and small for gestational age) and confounders. Additionally, variables that were highly correlated with gestational age were also included, namely maternal age, body mass index, parity, anaemia and birthweight. The ultrasound based gestational age, which was measured in women determined to be <24 weeks gestation by public sector midwives using last menstrual period (LMP) and/or measurement of the symphysis fundal height, was considered the gold standard measure. The variance of measurement error in the LMP based measure was estimated using the ultrasound based gestational age, with proxy values for the "true" value assigned based on expected values of gestational age.

The multiple overimputation procedure creates multiple imputation datasets (we chose 5). Multiple overimputation and the combination of estimates from multiple data sets is implemented the R package Amelia II while further analysis was conducted using another R package, Zelig, which allows pooling of the estimates from the separate analyses on the multiple overimputed data sets following Rubin's rules.

References

1. Blackwell M, Honaker J, King G. A unified approach to measurement error and missing data: Details and extensions. *Sociol Methods Res*. 2017;46(3):342-369.
2. Blackwell M, Honaker J, King G. A unified approach to measurement error and missing data: overview and applications. *Sociol Methods Res*. 2017;46(3):303-341.
3. Bachl M, Scharkow M. Some suggestions on dealing with measurement error in linkage analyses. 2017.

Supplementary File 9.3.3 Birthweight Findings

The overall incidence of high birthweight (HBW) was 4%; infants of preconception ART initiators were less likely to be HBW than infants of HIV-negative and initiating ART during pregnancy women. Among pregnancy initiators, third trimester initiators had a higher proportion of HBW infants (7%) than first (3%) and second (2%) trimester initiators.

Overall, 13% of infants were low birthweight (LBW); slightly higher among WLHIV (14%) than HIV-negative (12%) women. In adjusted analyses, LBW risk was similar for infants of WLHIV (ARR 1.12, 95% CI 0.91 - 1.38). LBW risk did not differ by ART status. Infants of preconception initiators (ARR 1.18, 95% CI 0.94, 1.50) and those initiating during pregnancy (ARR 1.06, 95% CI 0.82 - 1.39) were at similar LBW risk compared to infants of HIV-negative women. Among pregnancy initiators, LBW was higher with initiation in second (16%) trimester than first (12%) or third trimester (7%). In adjusted analyses, first (ARR 0.92, 95% CI 0.57 - 1.53) and third trimester (ARR0.46, 95% CI 0.22 - 0.99) initiators were at decreased risk of an LBW infant than HIV-negative women, but second trimester initiators (ARR1.35, 95% CI 0.99 - 1.83) were at increased risk of LBW.

9.4 Supplementary materials for Chapter 6

Supplementary Table 9.4.1 Baseline description of women with at least 3 pre-labour BP measurements vs those with <3 BP measurements

	Total N=3866	**≥3 BP measures N=1538**	**<3 BP measures N=2328**	**P value****
Age, years (%)				<0.0001
<24	1247 (32)	429 (28)	818 (35)	
25-29	1080 (28)	454 (30)	626 (27)	
>30	1581 (40)	655 (43)	884 (38)	
Median (IQR)	28 (23-32)	28 (24-32)	27 (23-32)	
Height, cm (%)				0.160
≤155	1140 (29)	470 (31)	670 (29)	
156-161	1359 (35)	533 (35)	826 (35)	
≥162	1105 (29)	412 (27)	693 (30)	
Missing	*262 (7)*	*123 (8)*	*139 (6)*	
Median (IQR)	158 (154-163)	158 (154-162)	158 (154-163)	
Body Mass Index, kg/m^2 (%)				0.003
Underweight (<18.5)	26 (0.7)	9 (0.6)	17 (0.7)	
Normal (18.5-24.9)	722 (19)	246 (16)	476 (20)	
Overweight (25.0-29.9)	1015 (26)	386 (25)	629 (27)	
Moderately Obese (30.0-34.9)	897 (23)	364 (24)	533 (23)	
Severely Obese (>35.0)	914 (24)	397 (26)	517 (22)	
Missing	*292 (8)*	*136 (9)*	*156 (7)*	
Median (IQR)	30 (26-36)	31 (26-36)	29 (25-35)	
Gravidity (%)				<0.0001
1	953 (25)	291 (19)	662 (28)	
2	1343 (35)	575 (37)	768 (33)	
≥3	1560 (40)	666 (43)	894 (38)	
Missing	*10 (0.3)*	*6 (0.4)*	*4 (0.2)*	
Median (IQR)	2 (2-3)	2 (2-3)	2 (1-3)	
Parity (%)				<0.0001
0	1178 (30)	383 (25)	795 (34)	
1	1452 (38)	644 (42)	808 (35)	
≥2	1226 (33)	505 (33)	721 (31)	
Missing	*10 (0.3)*	*6 (0.4)*	*4 (0.2)*	
Median (IQR)	1 (0-2)	1 (1-2)	1 (0-2)	
Previous Preterm (%)				0.251
Yes	286 (7)	127 (8)	159 (7)	
Haemoglobin g/dl (%)				0.790
Normal (≥11.0)	1457 (38)	567 (37)	890 (38)	
Mild Anaemia (9-10.9)	1065 (28)	401 (26)	664 (29)	
Moderate Anaemia (7-8.9)	229 (6)	94 (6)	135 (6)	
Severe Anaemia (<7)	5 (0.1)	2 (0.1)	3 (0.1)	

	Total N=3866	≥3 BP measures N=1538	<3 BP measures N=2328	P value**
Missing	*1110 (29)*	*474 (33)*	*636 (27)*	
Gestational Age Assessment*				
Median (weeks) (IQR)	18 (13-24)	18 (13-24)	18 (12-25)	0.089
Booking Blood Pressure (mmHG) (%)				0.095
Normal (Sys <120 or Dia <80)	2478 (64)	989 (64)	1489 (64)	
Prehypertensive (Sys 120-139 and Dia 80-89)	1223 (32)	471 (31)	752 (32)	
Hypertensive (Sys >140 or Dia >90)	165 (4)	78 (5)	87 (4)	
Booking Mean arterial pressure (mmHG) (%)				0.768
Low (<65)	72 (2)	26 (2)	46 (2)	
Normal (65-99)	3572 (92)	1428 (93)	2144 (92)	
High (>100)	216 (5)	84 (5)	132 (6)	
Missing	*6 (0.2)*	*0*	*6 (0.3)*	

* Best available measure (USS, SFH, LMP)
** Chi squared test/ rank-sum test for the presence of statistically significant difference between groups.

Supplementary Table 9.4.2 Average posterior probabilities of joint trajectory class assignments

Classes	Model 1			Model 2		
	Number of subjects (%)	Average Posterior Probability (%)		Number of Subjects (%)	Average Posterior Probability (%)	
1	1118 (73)	94.23	0.998 (0.953-0.999)	1122 (73)	94.38	0.998 (0.954-0.999)
2	420 (27)	89.86	0.977 (0.842-0.999)	416 (27)	89.59	0.977 (0.840-0.999)

Classes	Model 1			Model 2		
	Number of subjects (%)	Average Posterior Probability (%)		Number of Subjects (%)	Average Posterior Probability (%)	
1	611 (40)	85.58	0.919 (0.741-0.988)	810 (53)	90.10	0.981 (0.863-0.999)
2	810 (52)	90.10	0.981 (0.863-0.999)	611 (40)	85.80	0.921 (0.745-0.989)
3	117 (8)	91.26	0.981 (0.875-0.999)	117 (7)	97.45	0.975 (0.846-0.999)

Classes	Model 1			Model 2		
	Number of subjects (%)	Average Posterior Probability (%)		Number of Subjects (%)	Average Posterior Probability (%)	
1	391 (26)	85.51	0.914 (0.749-0.991)	775 (50)	81.83	0.866 (0.722-0.951)
2	775 (50)	82.24	0.868 (0.725-0.952)	391 (26)	85.51	0.913 (0.749-0.991)
3	310 (20)	86.62	0.866 (0.703-0.981)	310 (20)	83.22	0.874 (0.705-0.983)
4	62 (4)	89.46	0.987 (0.832-0.999)	62 (4)	89.46	0.988 (0.832-0.999)

Classes	Model 1			Model 2		
	Number of subjects (%)	Average Posterior Probability (%)		Number of Subjects (%)	Average Posterior Probability (%)	
1	384 (25)	85.24	0.914 (0.749-0.990)	9 (0.2)	99.80	0.999 (0.999-1.000)
2	9 (1)	94.48	0.973 (0.947-0.999)	768 (50)	81.53	0.863 (0.716-0.948)
3	768 (50)	81.80	0.865 (0.722-0.950)	384 (25)	85.24	0.914 (0.749-0.990)
4	318 (20)	83.32	0.866 (0.714-0.976)	318 (20)	80.49	0.852 (0.671-0.960)
5	59 (4)	82.39	0.890 (0.672-0.974)	59 (4)	85.93	0.938 (0.683-0.997)

Supplementary Table 9.4.3 Model fit statistics of joint trajectory classes

Model	BIC (N=73074)	BIC (n=10276)	AIC	>5% per group
2 Classes	-97870.52	-97849.93	-97773.93	Yes
3 Classes	-94728.19	-94694.84	-94571.80	Yes
4 Classes	-93533.43	-93485.36	-93308.04	Yes
5 Classes	-93135.40	-93070.67	-92831.83	No

Supplementary Table 9.4.4 Joint Systolic and Diastolic blood

	Joint group membership combinations*			
	Model 1 (Systolic BP)	Model 2 (Diastolic BP)	Prevalence	Combination Present
1	1	1	0.0%	No
2	2	1	0.0%	No
3	3	1	0.0%	No
4	**4**	**1**	**1.4%**	**Yes**
5	5	1	0.0%	No
6	1	2	0.0%	No
7	2	2	0.0%	No
8	**3**	**2**	**44.3%**	**Yes**
9	4	2	0.0%	No
10	5	2	0.0%	No
11	**1**	**3**	**22.9%**	**Yes**
12	2	3	0.0%	No
13	3	3	0.0%	No
14	4	3	0.0%	No
15	5	3	0.0%	No
16	1	4	0.0%	No
17	2	4	0.0%	No
18	**3**	**4**	**0.2%**	**Yes**
19	**4**	**4**	**24.1%**	**Yes**
20	5	4	0.0%	No
21	1	5	0.0%	No
22	**2**	**5**	**1.1%**	**Yes**
23	3	5	0.0%	No
24	4	5	0.0%	No
25	**5**	**5**	**6.1%**	**Yes**

*trajectory groups for both outcomes (systolic and diastolic BP)

Supplementary Table 9.4.5 Adjusted associations between baseline predictors and joint

Trajectory Group (Ref cat: Normal)	Predictor	ARR (95% CI)	P-value
Low Normal	Living with HIV (Ref cat: HIV negative)	1.24 (0.94 - 1.64)	0.123
	Age (Ref cat: <24 years)		
	25-29	0.83 (0.57 - 1.12)	0.320
	>30	0.79 (0.53 - 1.17)	0.240
	Body Mass Index (Ref cat: normal)		
	Overweight	0.65 (0.45 - 0.92)	0.016
	Obese	0.32 (0.22 - 0.45)	<0.0001
	Gravidity (Ref cat: 1)		
	2	1.87 (1.24 - 2.83)	0.003
	>3	2.14 (1.35 - 3.39)	0.001
High Normal	Living HIV (Ref cat: HIV negative)	1.16 (0.87 - 1.56)	0.317
	Age (Ref cat: <24 years)		
	25-29	0.71 (0.47 - 1.09)	0.121
	>30	1.45 (0.95 - 2.22)	0.087
	Body Mass Index (Ref cat: normal)		
	Overweight	1.44 (0.87 - 2.40)	0.159
	Obese	2.14 (1.34 - 3.40)	0.001
	Gravidity (Ref cat: 1)		
	2	0.69 (0.46 - 1.06)	0.091
	>3	0.65 (0.41 - 1.03)	0.068
Abnormal	Living with HIV (Ref cat: HIV negative)	0.96 (0.54 - 1.68)	0.875
	Age (Ref cat: <24 years)		
	25-29	1.28 (0.60 - 2.75)	0.523
	>30	1.77 (0.77 - 4.04)	0.177
	Body Mass Index (Ref cat: normal)		
	Overweight	0.79 (0.29 - 2.20)	0.656
	Obese	2.10 (0.91 - 4.89)	0.084
	Gravidity (Ref cat: 1)		
	2	0.50 (0.24 - 1.05)	0.065
	>3	0.35 (0.15 - 0.81)	0.014

Supplementary Table 9.4.6 Baseline description of women with at least 3 pre-labour blood pressure readings according to trajectory group

	Total N=1538	**Group 1: Normal N=767**	**Group 2: Low Normal N=384**	**Group 3: High Normal N=316**	**Group 4: Abnormal N=71**	**P-value**
Age, years (%)						0.002
<24	429 (28)	219 (29)	113 (29)	80 (25)	17 (24)	
25-29	454 (30)	246 (32)	119 (31)	69 (22)	20 (28)	
>30	655 (43)	302 (39)	152 (40)	167 (53)	34 (48)	
Median (IQR)	28 (24-32)	28 (24-32)	28 (23-32)	30 (24-34)	29 (25-34)	
Height, cm (%)						0.023
≤155	470 (30)	222 (29)	139 (36)	85 (27)	24 (34)	
156-161	533 (35)	277 (36)	134 (35)	102 (32)	20 (28)	
≥162	412 (27)	210 (27)	83 (22)	100 (32)	19 (27)	
Missing	*123 (8)*	*58 (8)*	*28 (7)*	*29 (9)*	*8 (11)*	
Median (IQR)	158 (154-162)	158 (154-162)	157 (153-161)	159 (155-163)	157 (153-163)	
Body Mass Index, kg/m^2 (%)						<0.0001
Underweight (<18.5)	9 (0.6)	4 (0.5)	3 (0.8)	2 (0.6)	0	
Normal (18.5-24.9)	246 (16)	110 (14)	104 (27)	25 (8)	7 (10)	
Overweight (25.0-29.9)	386 (25)	192 (25)	120 (31)	65 (21)	9 (13)	
Moderately Obese (30.0-34.9)	364 (24)	177 (23)	83 (22)	86 (27)	18 (25)	
Severely Obese (>35.0)	397 (26)	218 (28)	42 (11)	109 (34)	28 (39)	
Missing	*136 (9)*	*66 (9)*	*32 (8)*	*29 (9)*	*9 (13)*	
Median (IQR)	31 (26-36)	31 (27-37)	28 (24-32)	33 (29-38)	34 (30-40)	
Gravidity (%)						0.019
1	291 (19)	151 (20)	53 (14)	67 (21)	20 (28)	
2	575 (37)	293 (38)	157 (41)	102 (32)	23 (32)	
≥3	666 (43)	320 (42)	173 (45)	145 (46)	28 (39)	
Missing	*6 (0.4)*	*1 (0.1)*	*5 (0.9)*	*2 (0.6)*	*0*	
Median (IQR)	2 (2-3)	2 (2-3)	2 (2-3)	2 (2-3)	2 (1-3)	
Parity (%)						0.028
0	383 (25)	199 (26)	75 (20)	85 (27)	24 (34)	
1	644 (42)	330 (43)	173 (45)	116 (37)	25 (35)	
≥2	505 (33)	235 (31)	135 (35)	113 (36)	22 (31)	
Missing	*6 (0.4)*	*1 (0.1)*	*5 (0.9)*	*2 (0.6)*	*0*	
Median (IQR)	1 (1-2)	1 (0-2)	1 (1-2)	1 (0-2)	1 (0-2)	
Previous Preterm (%)						0.030
Yes	127 (8)	51 (7)	31 (8)	35 (11)	10 (14)	
Haemoglobin g/dl (%)						0.105
Normal (≥11.0)	567 (37)	297 (39)	126 (33)	116 (37)	28 (39)	
Mild Anaemia (9-10.9)	401 (26)	196 (26)	112 (29)	76 (24)	17 (24)	
Moderate Anaemia (7-8.9)	94 (6)	39 (5)	36 (9)	15 (5)	4 (6)	
Severe Anaemia (<7)	2 (0.1)	2 (0.3)	0	0	0	
Missing	*474 (32)*	*233 (30)*	*110 (29)*	*109 (34)*	*22 (31)*	
HIV/ART status (%)						0.204
HIV negative	967 (63)	494 (64)	226 (59)	197 (62)	50 (70)	
Living with HIV, preconception ART	306 (20)	149 (20)	83 (22)	67 (21)	7 (10)	
Living with HIV, initiating ART	265 (17)	124 (16)	75 (19)	52 (17)	14 (20)	
Gestational Age Assessment*						0.0344

	Total N=1538	Group 1: Normal N=767	Group 2: Low Normal N=384	Group 3: High Normal N=316	Group 4: Abnormal N=71	P-value
Median (weeks) (IQR)	18 (13-24)	18 (13-23)	18 (12-24)	20 (13-25)	17 (11-22)	
Booking Blood Pressure (mmHG) (%)						<0.0001
Normal (Sys <120 and Dia <80)	989 (64)	517 (67)	338 (88)	112 (35)	22 (31)	
Prehypertensive (Sys 120-139 and Dia 80-89)	471 (31)	228 (30)	44 (11)	169 (53)	30 (42)	
Hypertensive (Sys >140 or Dia >90)	78 (5)	22 (3)	2 (0.5)	35 (11)	19 (27)	
Booking Mean arterial pressure (mmHG) (%)						<0.0001
Low (<65)	26 (2)	10 (1)	15 (4)	1 (0.3)	0	
Normal (65-99)	1428 (93)	732 (95)	369 (96)	279 (88)	48 (68)	
High (>100)	84 (5)	48 (3)	0	36 (11)	23 (32)	

* Best available measure (USS, SFH, LMP)

9.5 Supplemental materials for Chapter 7

Supplementary Table 9.5.1 Adjusted associations between timing of ART initiation and preterm delivery among women with live singleton deliveries all 3 data sources

Data source	Outcome measure	Comparison C† (Ref category: ART Initiated before pregnancy)	ARR [95% CI]	P value
PHDC	Preterm delivery (<37 weeks)*	ART initiated during pregnancy	0.90 (0.85–0.95) [X]	<0.0001
			0.93 (0.88-0.99) [XX]	0.021
PER	Preterm delivery (<37 weeks)**	ART initiated during pregnancy	1.09 (0.91-1.32) [X]	0.356
			1.12 (0.91-1.38) [XX]	0.297
Cohort	Preterm delivery (<37 weeks)***	ART initiated during pregnancy	0.84 (0.45-1.58) [X]	0.841
			1.08 (0.53-2.18) [XX]	0.597

ART – antiretroviral therapy, PHDC – Provincial health data center, PER – pregnancy exposure registry, ARR – adjusted risk ratio

* adjusted for age, gravidity, TB, smoking
** adjusted for age, gravidity, wasting
*** adjusted for age, gravidity, BMI, hypertension, hazardous drinking, depressive symptoms, PI regimen

Supplementary Table 9.5.2 Baseline comparisons of common variables across the three data sources

	PHDC N=183 593		PER N=9476		Cohort Study N=989	
	HIV-uninfected N=148 593	**HIV-infected N=35 353**	**HIV-uninfected N=6473**	**HIV-infected N=3003**	**HIV-uninfected N=510**	**HIV-infected N=479**
Age, years						
<24	59549 (40)	7861 (22)	2690 (42)	644 (21)	185 (36)	82 (17)
25-29	42 119 (28)	10527 (30)	1848 (29)	893 (30)	161 (32)	133 (28)
>30	46 925 (32)	16965 (48)	1935 (30)	1466 (49)	164 (32)	264 (55)
Median	26 (22-31)	29 (25-33)	26 (22-31)	29 (25-34)	26 (23-32)	30 (26-35)
Gravidity						
1	81 144 (55)	16 781 (48)	1904 (29)	497 (17)	136 (27)	69 (14)
2	44 644 (30)	11 732 (33)	2051 (32)	988 (33)	161 (32)	126 (26)
≥3	22 804 (15)	6840 (19)	2486 (38)	1487 (50)	213 (42)	284 (59)
Missing	*1 (0)*	*0*	*32 (1)*	*31 (1)*	*0*	*0*
Median	*1 (1-2)*	*2 (1-2)*	*2 (1-3)*	*2 (2-3)*	*2 (1-3)*	*3 (2-3)*
Parity						
0	90 086 (61)	19 249 (54)	1383 (21)	531 (18)	168 (33)	83 (17)
1	38 767 (26)	10 284 (29)	2172 (34)	1084 (36)	169 (33)	162 (34)
≥2	19 740 (13)	5820 16)	2842 (41)	1346 (45)	173 (34)	234 (49)
Median	0 (0-1)	0 (0-1)	*76 (1)*	*43 (1)*	2 (1-3)	1 (1-2)
HIV status	161 634 (82%)	35 879 (18%)	6961 (70%)	3049 (30%)	510 (52%)	479 (48%)
ART status						
No ART initiated	-	1 668 (5%)	-	0	-	0
Initiated before pregnancy	-	19 966 (55%)	-	2014 (66%)	-	337 (70%)
Initiated during pregnancy	-	14 245 (40%)	-	1035 (34%)	-	142 (30%)

ART – antiretroviral therapy, PHDC – Provincial health data center, PER – pregnancy exposure registry

Supplementary Table 9.5.3 Adjusted associations between HIV/ART exposure groups and preterm delivery among women with live singleton deliveries in all three data sources adjusted by common variables (age and gravidity)

Data source	Outcome measure	Comparison A[†] (Ref category: HIV-uninfected)			Comparison B[††] (Ref category: HIV-uninfected)		
			ARR [95% CI]	P value		ARR [95% CI]	P value
PHDC	Preterm delivery (<37 weeks)	HIV-infected	1.16 (1.12–1.19)	<0.0001	No ART initiated	1.53 (1.37-1.72)	<0.0001
					ART initiated before pregnancy	1.17 (1.12-1.21)	<0.0001
					ART initiated during pregnancy	1.10 (1.05-1.16[)]	<0.0001
PER	Preterm delivery (<37 weeks)	HIV-infected	1.03 (0.92-1.15)	0.656	ART initiated before pregnancy	0.98 (0.86-1.12) [X]	0.734
					ART initiated during pregnancy	1.12 (0.95-1.32) [X]	0.168
Cohort	Preterm delivery (<37 weeks)	HIV-infected	1.03 (0.69-1.55)	0.869	ART initiated before pregnancy	1.04 (0.67-1.63)	0.848
					ART initiated during pregnancy	1.01 (0.55-1.87)	0.973

ART – antiretroviral therapy, PHDC – Provincial health data center, PER – pregnancy exposure registry, ARR – adjusted risk ratio

9.6 PIMS Cohort Profile

Malaba TR, Gray CM, Myer L, Newell ML, for the PIMS Study Group. Cohort profile: Prematurity Immunology in HIV-infected Mothers and their infants Study (PIMS).
BMJ Open. 2021;11(9): e047133.

Relevance of this paper to the thesis

This chapter presents data on the impact of different gestational age assessment methods on the incidence of preterm delivery and small for gestational age infants. Additionally, there is exploration of the impact of these methods on the association between maternal HIV and antiretroviral therapy use and adverse birth outcomes. The use of different assessment methods has previously been hypothesised as a methodological issue that could explain the heterogeneous findings in previous studies.

Contribution of the student and co-authors

TRM conceptualised and conducted all analyses, led data interpretation and drafted the manuscript. TRM directed data collection, with assistance from HM and AP. MLN and LM conceptualised the study from which these data arise, were responsible for funding, implementation and overall leadership, and assisted with data interpretation. All authors reviewed and approved the final manuscript.

TRM, LM, CG and MLN: Conceptualization and design of the study

TRM: Study conduct, data collection and statistical analysis.

TRM and MLN: Interpretation of data and writing the manuscript

TRM, LM, CG and MLN review and editing

9.6.1 Abstract

9.6.1.1 Purpose

PIMS is a prospective cohort study in South Africa investigating the association between antiretroviral therapy (ART) use, preterm delivery (PTD) and small-for gestational age (SGA) live births. PIMS main hypotheses are that ART initiation in pregnancy and ART-induced hypertension are associated with PTD and SGA respectively and that reconstitution of cellular immune responses in women on ART from before pregnancy results in increases in PTD of appropriate-for-gestational age (AGA) infants.

9.6.1.2 Participants

Pregnant women (n=3972) aged ≥18 years regardless of HIV status recruited from 2015 to 2016 into the overall PIMS cohort (2517 living without HIV, 1455 living with HIV). A nested cohort contained 551 women living with HIV who were ≤24 weeks' GA on ultrasound: 261 initiated ART before pregnancy, 290 initiated during pregnancy.

9.6.1.3 Findings to date

Women in the overall cohort were followed antenatally through to delivery using routine clinical records; further women in the nested cohort were actively followed up until 12 months postpartum, with data collected on maternal health (HIV care and ART use, clinical care and inter-current clinical history). Other procedures conducted on the nested cohort included physical examinations (anthropometry, blood pressure measurement), assessment of fetal growth (ultrasound), maternal and infant phlebotomy for storage of plasma, RNA and peripheral blood mononuclear cells, collection of delivery specimens (placenta and cord blood), and infant 12 month developmental assessment. Preliminary findings have contributed to our understanding of risk factors for adverse birth outcomes, and the relationship between pregnancy immunology, HIV/ART and adverse birth outcomes.

9.6.1.4 Future plans

Using specimens collected from study participants living with HIV throughout pregnancy and first year of life, the PIMS provides a valuable platform for answering a variety of research questions focused on temporal changes of immunology markers in women whose immune status is altered by HIV infection, and how ART initiated during pregnancy affects immune responses. The relationship between these immunological changes with adverse birth outcomes as well as possible longer-term impact of exposure to ART in fetal and early life will be explored. Additionally, further active, and passive follow-up of mothers and their infants is planned at school-going age and beyond to chart growth, morbidity, and development, as well as changes in family circumstances.

9.6.2 Introduction

Antiretroviral drugs in pregnant women living with HIV prevent mother-to-child transmission (PMTCT) and delay HIV disease progression. WHO guidelines now recommend antiretroviral therapy (ART) for all, immediately upon HIV diagnosis, including for women living with HIV during pregnancy and breastfeeding, to be continued lifelong (1). However infants born to mothers living with HIV would be exposed to multi-drug ART regimens for prolonged periods at a crucial time during their development (2), which could result in decreased health, developmental, and survival outcomes (3, 4). In high maternal HIV prevalence settings the increasing population of ART-exposed infants could make the goal of under-five mortality reduction less likely.

Adverse birth outcomes contribute significantly to under-five mortality, as well as infant health and developmental problems (5). There is an ongoing debate regarding the association between exposure to maternal ART during fetal life and adverse birth outcomes, following reports from Europe (6-9), USA (10) and Africa (11-14) of possible ART-associated increased risk of preterm delivery (PTD), small-for-gestational age (SGA) or low birthweight (LBW) infants. Furthermore, these exposed infants are also at increased risk of acquiring viral infections (15, 16), as well as the negative impact of ART on fetal brain development and function (17). Interpretation of findings from various studies, especially from African settings, is hindered by the reliability of gestational assessment, with ultrasound dating in early pregnancy usually unavailable.

There is limited understanding of general pregnancy-related immune changes in high HIV prevalence settings. Successful pregnancies require intricate fetal-maternal (FM) immune balances, to enable maternal tolerance of the semi-allogeneic fetus; this FM tolerance is primarily maintained by the placenta (18, 19). Consequently, adverse birth outcomes could be hypothesized to be due to placental interface FM tolerance disruption because of cytokine shifts associated with ART initiation causing early initiation of uterine contractions (20). Additionally, there are suggestions that adverse birth outcomes could also be associated with ART-induced dysregulation of maternal and infant metabolism. In order to meet the specific needs during pregnancy and infancy the metabolism is tightly regulated, however ART is known to interfere with lipid metabolism (21). The emerging field of immune metabolism has shown that alterations in the lipid profile increases susceptibility to viral infections by skewing immune responses. The complex interplay between pregnancy, HIV/ART, host immunity, adverse birth outcomes and long-term child health is poorly understood as detailed data are few, and related to drug combinations which may no longer be in use. Further research is

required to examine epidemiologic and immunological associations and inform understandings of underlying biological mechanisms.

An increase in PTD rates, especially of SGA infants, could impact on the long-term growth and development of children, and would have consequences for the health and wellbeing of their families and population more widely. We therefore aimed to improve understanding of maternal immune profiles during pregnancy in the context of ART use during gestation, adverse birth outcomes and long-term child health in Cape Town, South Africa, an area of high HIV prevalence. This manuscript presents the details of the setting up of the cohort, including aims and objectives and a description of baseline findings along with other preliminary findings.

9.6.3 Aim and Objectives

The primary focus of the PIMS study was to investigate and quantify the risk of preterm and SGA deliveries; with underlying hypotheses that (i) timing of ART use (from before or during pregnancy) is associated with increased risk of PTD, (ii) ART-induced hypertension during pregnancy results in increased risk of SGA and (iii) reconstitution of cellular immune responses during ART in pregnancy results in increases in PTD of appropriate-for-gestational age (AGA) infants. Our secondary focus was to determine long-term (first five years of life) child health outcomes of PTD infants (by weight at gestational age and maternal HIV/ART status). Our hypotheses are that (i) throughout childhood SGA infants are disadvantaged in terms of growth and development compared to preterm AGA infants and (ii) ART use alters maternal and fetal lipid metabolism resulting in susceptibility to infections and alterations in vaccine responses in childhood.

9.6.4 Cohort Description

9.6.4.1 Setting

Between April 2015 and October 2016, we enrolled pregnant women (aged ≥18 years) at their first antenatal care (ANC) visit in a prospective cohort study, at a single large public sector primary care facility (Gugulethu Midwife Obstetric Unit (MOU)) in a low-income high HIV-prevalence sub-district of Cape Town, South Africa.

9.6.4.2 Study Design

The overall prospective, observational design includes two 'nested' groups of pregnant women:

- **Group 1:** the overall population of pregnant women (≥18 years) seeking ANC services at Gugulethu MOU during a 18-month period; within this group, a subset of women thought to be ≤24 weeks' gestation based on history or examination (clinical GA) underwent ultrasound scan by a research sonographer for more accurate gestation estimation; enrolled into observation group.
- **Group 2:** all pregnant WLHIV seeking ANC who are ≤24 weeks' gestation at US at their first ANC visit, regardless of current ART use at the first ANC visit (nested within Group 1). Enrolled into longitudinal cohort with data collection through questionnaires, clinical assessments and phlebotomy spanning pregnancy to early infancy.

This study design enables quantification of the risk of adverse birth outcomes in the overall cohort, as well as the more detailed Group 2 group also enabling investigation of the consequences of the immune response following ART initiation in pregnancy for onset of labour and preterm delivery.

9.6.4.3 Ethical Approval

The study was reviewed and approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (UCT HREC 739/2014) and the University of Southampton Faculty of Medicine Ethics Committee (12542 PIMS).

9.6.4.4 Routine Care Services

As part of routine ANC services, gestational age (GA) was estimated based on date of last menstrual period (LMP) and symphysis-fundal height (SFH) by public sector midwives. All women without a previous HIV diagnosis underwent HIV testing, with universal ART eligibility. Women living with HIV conceiving while on ART continued their current regimen throughout pregnancy; regimens included NNRTIs such as efavirenz (EFV) or protease inhibitor (PI, predominantly used after failure of first-line therapy). For women initiating ART in pregnancy, a fixed-dose combination of tenofovir (TDF) + emtricitabine (FTC) + efavirenz (EFV) was used throughout.

9.6.4.5 Recruitment

Following screening of all women attending their first ANC visit, those ≥18 years were eligible and approached to participate in the study. Following screening, ineligible women were referred back to their ANC clinics in line with the Western Cape Department of Health's health care model. Women

who agreed to participate had their routinely collected LMP-based GA and SFH-based GA reviewed by the counsellor; women estimated to be ≤24 weeks were referred for a research ultrasound scan (US) for formal pregnancy dating by a research sonographer blinded to the midwife assessment. Women living with HIV who were ≤24 weeks' gestation on US were then recruited into a nested cohort (Group 2); half of these had initiated life-long ART prior to conception and half initiated ART during pregnancy.

All participants provided written informed consent prior to study participation, with re-consenting of mother-infant pairs at the first postpartum visit for paediatric follow-up. Consent for study participation included data abstraction from routine clinical records through the pregnancy and post-partum period.

9.6.4.6 Participant Baseline Characteristics

A total of 4431 women registered for ANC during the study recruitment period, of whom 4111 (93%) were screened for the study and 3972 (90%) enrolled; all delivered by May 2017 (Figure 9.6.1). Main reasons for being screened out were under-age, referrals from Basic Antenatal Clinics (BANC) or not being interested. Of the enrolled women, 2517 (63.4%) were living without HIVand 1455 (36.6 %) WLHIV (Table 9.6.1); 2199 (55.4%) were referred for ultrasound based on their clinical GA, 1327 (60.3%) were living without HIV and 872 (39.7%) living with HIV.

Median age at enrolment was 28 years (IQR 23-32), with women living without HIV younger than WLHIV women, and those initiating ART preconception older than WLHIV initiating ART during pregnancy. In line with age differences, WLHIV were of higher gravidity than women living without HIV, but the difference in parity was small; women who initiated ART pre-pregnancy were of a higher gravidity than those initiating during pregnancy. A quarter (25%) of all women were overweight, while over half (55%) were obese, with little or no difference between groups. Having previously had a preterm delivery was more common among WLHIV than women living without HIV. Mild anaemia was relatively common in all groups, especially in women initiating ART during pregnancy (Table 9.6.1). Overall, 3479 (87.6%) women had gestation estimated by LMP, 2327 (58.6%) by SFH and 2334 (58.8%) by ultrasound; with estimated median GA at enrolment visit varying by assessment method (Table 9.6.1).

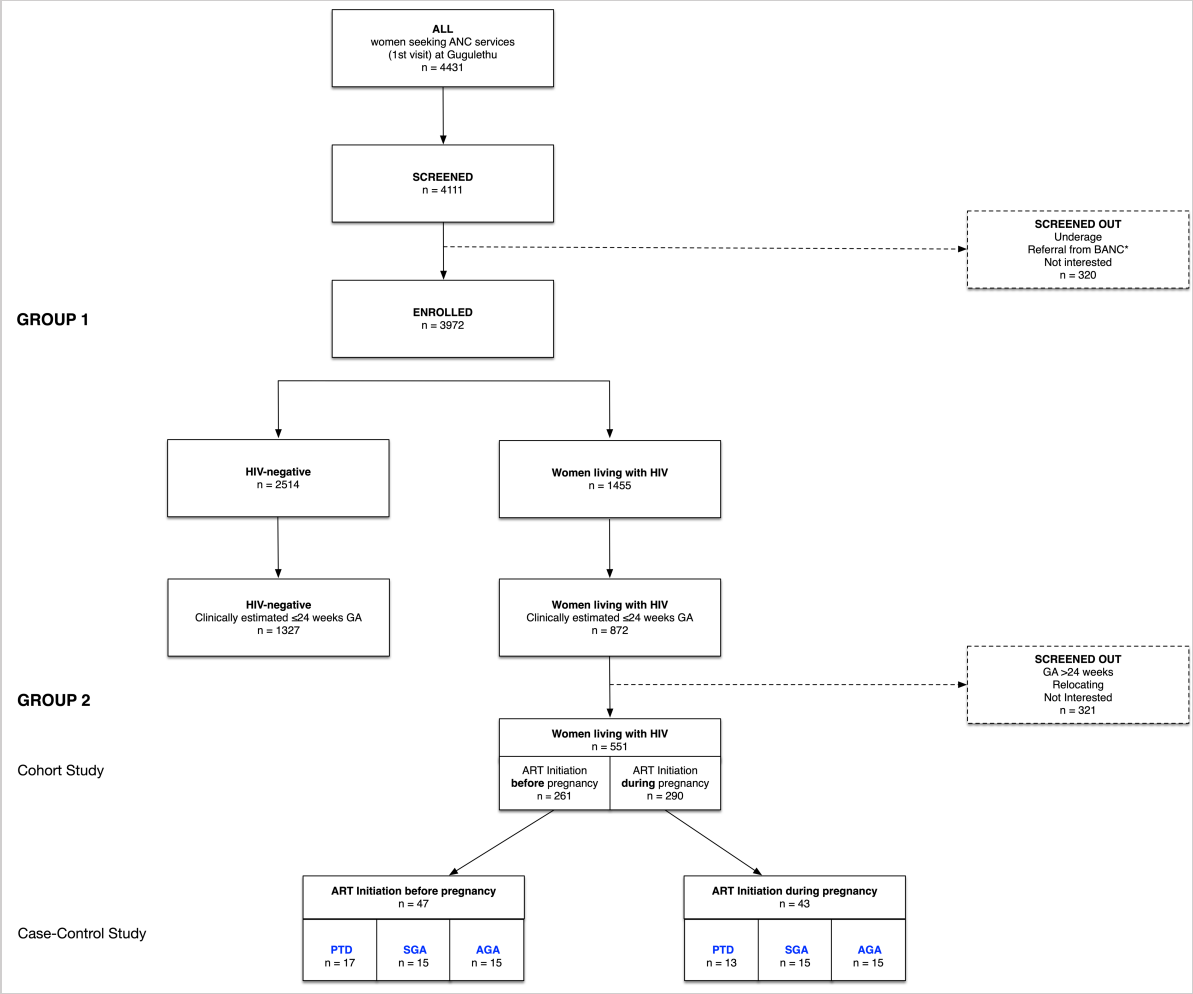

Figure 9.6.1 Cohort profile

AGA - appropriate-for-gestational age, ART - antiretroviral therapy, GA - gestational age, PTD – preterm delivery, SGA - small-for-gestational age
* BANC – basic antenatal clinic

There were 1455 WLHIV in Group 1; 718 (49.3%) were ≤24 weeks on ultrasound, of whom 551 were enrolled into Group 2 (Figure 9.6.1). The likelihood of inclusion into Group 2 did not differ by baseline characteristics (Table 9.6.2). In comparison to Group 1 participants, Group 2 participants differed in age, gravidity, parity and previous PTD, likely driven by the women living without HIV (Table 9.6.3). In multivariable regression allowing for HIV infection, the only difference that persisted between these groups was age (Table 9.6.4). When Group 2 women were compared to other WLHIV (not enrolled in Group 2), they were slightly older, more likely to have normal haemoglobin levels (≥11g/dl) and lower gestational age (Table 9.6.3). In multivariable regression the only difference between these groups that persisted was gestational age, which was a Group 2 inclusion criterion (Table 9.6.4).

Of the 551 WLHIV in Group 2, 261 (47%) initiated before pregnancy and 290 (53%) initiated during pregnancy (Table 9.6.5). Women who initiated ART during pregnancy were on average younger, and of lower gravidity. Overall, three-quarters of Group 2 women were overweight or obese, with little difference by ART status. Of the women who initiated ART during pregnancy, the majority (64%) had tested HIV positive in the index pregnancy; the rest had previously tested positive although were not on ART at conception. In line with local and WHO treatment guidelines, most women (91%) were on a regimen of two NRTIs (TDF +FTC), plus NNRTI EFV. PI usage in this cohort was low, at only 9% of women who initiated pre-pregnancy and 1% of women who initiated during pregnancy. Median CD4 count was 433 cells/µl overall, 527 cells/µl in women who initiated ART before pregnancy and 373 cells/µl in women initiating at the first antenatal visit. There were no differences in smoking or drug usage between these two groups of women, but women who initiated during pregnancy were more likely to have ever consumed alcohol or consumed in the last 30 days (Table 9.6.5).

9.6.4.7 Study Follow-up

At baseline, all women (Group 1) had clinical and medical history, routine first ANC visit physical examination, screening tests and GA assessment data collected via abstraction of the Maternity Case Record (MCR) booklet (Table 9.6.5), which is a standardised patient-held maternity record used by all facilities providing maternity services to record clinical data from the antenatal through to postpartum period, including labour. The MCR also serves as a referral letter, thus serving as a link between antenatal and labour care. In addition, the National Health Laboratory Services (NHLS) database was searched for CD4, Viral load results and other laboratory values not recorded in the MCR. Further follow-up for women in Group 1, not eligible for Group 2, was through data abstraction of the MCR following discharge from the postnatal ward (MCR retained at delivery facility). Data was abstracted from follow-up ANC notes, clinical notes during labour and newborn assessments (Table 9.6.5).

Women in Group 2 participated in up to eight scheduled study visits, from the start of ANC through to 12 months postpartum. Women on ART from before pregnancy had three antenatal visits at <24 weeks, 28 and 34 weeks; women who initiated ART during pregnancy had an additional study visit two weeks after the ART initiation (which in most women took place on the same day, or close to, the first study visit). Following delivery, women were re-consented for infant participation, and study visits were conducted <7days, 10 weeks, 6 months and 12 months postpartum. At all study visits, data were collected on maternal health (HIV care and ART use, clinical care and inter-current

clinical history). Other procedures included physical examinations (anthropometry, blood pressure measurement), phlebotomy (50ml) for storage of plasma, peripheral blood mononuclear cells (PBMC) and RNA; a follow-up ultrasound was conducted at 28 weeks to assess fetal growth (Table 9.6.5).

At postpartum study visits, additional data was collected on infant health (including infant feeding and inter-current clinical history) and physical examination of infants was conducted (basic anthropometry). At the 12 month visit, developmental assessment was conducted using the Ages and Stages questionnaire (22) – a general developmental screening tool testing five key areas: personal social, gross motor, fine motor, problem solving and communication skills. Infant specimen collection included Dried Blood Spot (DBS) sampling and storage at 10 weeks study visit, and phlebotomy (2ml) for measurement of immunological functioning and antibody responses to routine childhood immunisations (rotavirus and measles) at 12 month study visit. In addition, data on infant health status, including vaccinations, chemoprophylaxis use (including nevirapine and co-trimoxazole) and routine HIV PCR testing, was abstracted from Road-to-Health Booklets - patient-held booklet taken to all clinical and immunisation visits used to monitor infant growth and development until age 5 years (Table 9.6.6).

Further active follow-up of the women and their infants will occur at school-going age and beyond to chart growth, morbidity and development, as well as changes in family circumstances. It is envisaged that subsequent to this, longer term follow-up will be passive through the use of routinely collected data. The Western Cape Provincial Department of Health's public-sector patient administration systems all share unique health identifier (23); data relating to participants health service contacts, health conditions and health outcomes for specific conditions will be obtained from the Provincial Health Data Centre, which consolidates person-level clinical data across government services (24).

9.6.4.8 Data Collection

An overview of the main included data collection instruments is presented in Table 9.6.6, covering self-reported information on clinical history, ART use and adherence, and medical events; as well as information obtained from routinely collected data.

9.6.4.9 Specimen Collection

To investigate the proposed hypothesis that immunological changes resulting from maternal ART exposure are associated with adverse birth outcomes, women enrolled into the follow-up cohort were intensively sampled, with repeated phlebotomy throughout pregnancy and the postpartum period for immunological investigations (Table 9.6.7). Using samples from all antenatal plasma, inflammatory markers (C-reactive Protein, Serum Amyloid A and CCL10 (IP-10) in women are being measured. Further, following a nested case-control design in Group 2 (n=90), investigations compare women who delivered preterm (PTD cases) or had small-for-gestational age infants (SGA cases) and those from appropriate controls (term AGA) (matched for GA and ART status) (Figure 9.6.1). Investigations include longitudinal quantification of plasma cytokine profiles, phenotypic and functional characterisation of regulatory T cells (Tregs), antigen-presenting cells and metabolites associated with mitochondrial functioning and lysophospholipids (Figure 9.6.2). The combined studies of these immune parameters will inform understanding of ART use during pregnancy on the areas of the immune system that have been shown to be critically involved in regulating maternal immune tolerance to the fetus, and their associations with onset of labour and preterm delivery.

At delivery, placentas and cord bloods were collected whenever possible, a scoring algorithm was developed which graded placentas and dictated specimen processing according to membrane completeness and time received in laboratory relative to delivery time (Table 9.6.8). Using flow cytometry and tissue immunostaining techniques the following investigations will be conducted: examination of the effect of HIV infection/ART exposure on the phenotypic characteristics and functionality of placental macrophages (Hofbauer cells and decidual macrophages) at the maternal-fetal interface and placental Tregs and their association with adverse birth outcomes. Characterization of cord blood Tregs and correlation of their frequency, function, and phenotype with placental Tregs and birth outcomes (Figure 9.6.2).

9.6.4.10 Study Retention

Loss to follow-up was categorised based on the last visit before the woman was lost; in total 158 (29%) women were lost to follow-up (LTFU) (Figure 9.6.3). Women lost before the post-delivery

study visit (n=88), either experienced pregnancy losses (n=37), were no longer interested in participating (n=24) or relocated out of the study area (n=17). For women LTFU between delivery and the 6 month postpartum visit (n=24), reasons included relocation (n=8), not contactable (n=6), no longer interested in participating (n=5) and maternal/infant death (n=5) (Figure 9.6.2). For women LTFU between 6 month and 12 month postpartum (n=46), reasons included relocation (n=15), not contactable (n=26), no longer interested in participating (n=2) and maternal/infant death (n=3).

There were no appreciable differences by ART status in women LTFU before post-delivery visit (RR 0.81, 95% CI 0.55 – 1.19). However, women who initiated ART before pregnancy were less likely to be LTFU between delivery and 6 months postpartum (RR0.44, 95% 0.19 – 1.04) and between 6 month and 12 month postpartum (RR0.58, 95% 0.33 – 1.02). No baseline characteristics were associated with LTFU.

9.6.5 Findings to date

9.6.5.1 Gestational Age Assessment

In the overall cohort, 1787 women with live singleton births were included in the analysis of the association between HIV status and timing of ART initiation and PTD by GA assessment method used (last menstrual period (LMP), measurement of symphysis fundal height (SFH) and ultrasound (US). Using US-GA, PTD risk was associated with maternal HIV infection and ART use, with WLHIV, on ART from before or early pregnancy, almost twice as likely to deliver preterm than women living without HIV (25). A weaker association was observed when GA assessment was based on SFH, while with LMP-GA the difference by HIV status was minimal. We did not find any appreciable differences in the PTD risk for WLHIV by timing of ART initiation across all three assessment methods (25). Our findings (in both the overall cohort and in women with all three assessments) suggest that methods of GA assessment explain at least partially the heterogeneity of findings from previous studies on the association between ART use and adverse birth outcomes, suggesting that care should be taken when interpreting results from such studies.

9.6.5.2 Obesity

In the overall cohort, 2921 women with live singleton births were included in the analysis of the association between maternal body mass index and adverse birth outcomes. In a subset cohort the association between gestational weight gain (GWG) and adverse birth outcomes was examined. Maternal obesity was associated with increased likelihood of having high birthweight and large size

for gestational age infants. In the subset cohort, GWG was associated with increased likelihood of spontaneous PTD and high birth weight infants (26). Obesity during pregnancy is prevalent in this setting and appears associated with increased risk of adverse birth outcomes in both WLHIV and women living without HIV.

9.6.5.3 Placental Pathology

Preliminary analysis of placental histopathology from a sub-set of women enrolled in the prospective cohort showed significant associations between placental pathology and adverse birth outcomes: presence of focal infarction was associated with increased risk of low birthweight (LBW); the lower the weight of the basal plate weight led to increased risk of LBW, PTD and SGA; and prolonged meconium exposure was associated with increased risk of SGA. These findings suggest that adverse birth outcomes are driven primarily by placental abnormalities which do not appear to be associated with the ART initiation timing (27). Immunofluorescence and immunohistochemistry staining were performed on these wax blocks to identify regulatory T cells along with macrophages. Further analysis is ongoing.

Within the placenta, investigation of the distribution of pro-inflammatory (M1) and anti-inflammatory (M2) placental macrophages at the maternal-fetal interface showed no differences in the tissue density of these macrophages within the decidual membranes and villous tissue according to timing of ART initiation. Data suggest that the Hofbauer Cells (which are fetal macrophages) are not polarized into M1/M2 phenotypes but are rather “intermediate” types (28).

9.6.6 Strengths and limitations

Key strengths of the PIMS study include the recruitment of a large community-based cohort in an area of high HIV prevalence; the observational nature of the study provides good external validity of experiences of a public sector primary care population over pregnancy. A further strength lies in the use of a research sonographer for the GA assessment in women ≤24 weeks when US is highly reproducible and accurate (while routinely used clinical assessments are less reliable) which is particularly important when studying associations with adverse birth outcomes, as compromises in outcome ascertainment methods can affect the detection of the magnitude of associations.

Additionally, the maternal biological specimen from Group 2 at three or four times (depending on ART status group) throughout pregnancy and at delivery is an important strength because it enables immunological, metabolomic and placental investigations to inform understanding of mechanisms

underlying adverse birth outcomes in in women living with HIV. As pregnancy is a state of immunoregulation requiring tolerance of a semi-allogeneic fetus, the assessment of placentas of enrolled women provides a unique opportunity to investigate the link between HIV, ART and adverse birth outcomes. Collection of infant specimens further strengthens the study as it is one of the first studies combining metabolomic and immunologic assessments. This will provide an integrated model of the immune-metabolism association in HIV-exposed infants and the consequences of maternal metabolic dysregulation for the immune responses of the infant. Furthermore, the developmental assessments carried out in the infants provides the opportunity to consider the association between maternal immune function during pregnancy and early childhood immunological and developmental outcomes.

PIMS has valuable sub-designs in addition to the observational study design with the overall cohort stratified by maternal HIV status, ART use, and related risk factors, with all details to be analysed by timing of ART initiation in line with the main PIMS objectives. One sub-design is the cohort study in women living with HIV, who have data collection through questionnaires, clinical assessments and phlebotomy spanning pregnancy to early infancy. This study design enables quantification of the risk of adverse birth outcomes in the overall cohort, as well as the more detailed Group 2 group also enabling investigation of the consequences of the immune response following ART initiation in pregnancy for onset of labour and preterm delivery. Another sub-design is the nested case–control study which will enable immunological investigations in women who did and did not delivery preterm/SGA infant. The ability to track patients using the Western Cape unique identifier across different health and laboratory services, enables the passive long term follow-up of the Group 2 women and their infants; with available data including patient level data (administrative, demographic and clinical), visit level data (clinical observations and findings), laboratory tests and medication.

A limitation of the study is while women living without HIV are included in Group 1, providing a comparison group for birth outcome and various maternal characteristics, Group 2 did not include living without HIV or ART-unexposed comparator groups. As such the detailed immunological analyses over pregnancy are limited to WLHIV, with timing of ART initiation a main explanatory variable. Additionally, we do not have full detailed information on all Group 1 women and were instead limited by routinely collected data in medical notes some information relating to maternal characteristics was based on self-report and thus subject to potential biases. In order to address this limitation, routinely collected data was also collected to confirm data on birth outcomes.

9.6.7 Findings to date

Using data collected from study participants living with HIV throughout pregnancy and first year of life, PIMS provides a valuable platform for answering a variety of research questions related to maternal and child health. In particular, PIMS well equipped to investigate temporal changes of immunology markers in women whose immune status is altered by HIV infection, and how ART initiated during pregnancy affects immune responses. The relationship between these immunological changes with adverse birth outcomes as well as possible longer-term impact of exposure to ART in fetal and early life will be explored. Additionally, through use of the Western Cape Department of Health unique identifier further active and passive follow-up of mothers and their infants is planned at school-going age and beyond to chart growth, morbidity and development, as well as changes in family circumstances.

9.6.6 Acknowledgements

The authors are extremely grateful to all mothers who took part in this study. We acknowledge the Gugulethu Midwife Obstetric Unit staff and the whole research team including Hlengiwe Madlala, Colin Newell, Megan Mrubata and Lee-Ann Stemmet for their valuable input and contribution.

9.6.7 References

1. WHO. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. Geneva, Switzerland: World Health Organization; 2015.
2. UNAIDS. On the fast-track to an AIDS-free generation. Geneva, Switzerland: United Nations Joint Program on HIV/AIDS; 2016.
3. Brennan AT, Bonawitz R, Gill CJ, et al. A meta-analysis assessing all-cause mortality in HIV-exposed uninfected compared with HIV-unexposed uninfected infants and children. *AIDS.* 2016;30(15):2351-2360.
4. le Roux SM, Abrams EJ, Nguyen K, Myer L. Clinical outcomes of HIV-exposed, HIV-uninfected children in sub-Saharan Africa. *Trop Med Int Health*. 2016;21(7):829-845.
5. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet.* 2015;385(9966):430-440.
6. Martin F, Taylor GP. Increased rates of pre-term delivery are associated with the initiation of highly active antiretroviral therapy during pregnancy: a single centre cohort study. *J Infect Dis*. 2007;196(4):558–561.
7. Sibiude J, Warszawski J, Tubiana R, et al. Premature delivery in HIV-infected women starting protease inhibitor therapy during pregnancy: role of the ritonavir boost? *Clin Infect Dis.* 2012;54(9):1348-1360.
8. Thorne C, Patel D, Newell M. Increased risk of adverse pregnancy outcomes in HIV-infected women treated with highly active antiretroviral therapy in Europe. *AIDS*. 2004;18(17):2337-2339.
9. Townsend CL, Cortina-Borja M, Peckham CS, Tookey PA. Antiretroviral therapy and premature delivery in diagnosed HIV-infected women in the United Kingdom and Ireland. *AIDS*. 2007;21(8):1019-1026.
10. 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):1195-1201.
11. Chen JY, Ribaudo HJ, Souda S, et al. Highly active antiretroviral therapy and adverse birth outcomes among HIV-infected women in Botswana. *J Infect Dis*. 2012;206(11):1695-1705.
12. Fowler MG, Qin M, Fiscus SA, et al. Benefits and risks of antiretroviral therapy for perinatal HIV prevention. *N Engl J Med*. 2016;375(18):1726-1737.
13. Li N, Sando MM, Spiegelman D, et al. Antiretroviral therapy in relation to birth outcomes among HIV-infected women: a cohort study. *J Infect Dis*. 2016;213(7):1057-1064.

14. Powis KM, Kitch D, Ogwu A, et al. Increased risk of preterm delivery among HIV-infected women randomized to protease versus nucleoside reverse transcriptase inhibitor-based HAART during pregnancy. *J Infect Dis*. 2011;204(4):506-514.
15. Slogrove AL, Cotton MF, Esser MM. Severe infections in HIV-exposed uninfected infants: clinical evidence of immunodeficiency. *J Trop Pediatr*. 2010;56(2):75-81.
16. Cotton MF, Slogrove AL, Rabie H. Infections in HIV-exposed uninfected children with focus on sub-Saharan Africa. *Pediatr Infect Dis J*. 2014;33(10):1085-1086.
17. Le Doaré K, Bland R, Newell ML. Neurodevelopment in children born to HIV-infected mothers by infection and treatment status. *Pediatrics*. 2012;130(5):e1326-e44.
18. Martínez-Varea A, Pellicer B, Perales-Marín A, Pellicer A. Relationship between maternal immunological response during pregnancy and onset of preeclampsia. *J Immunol Res*. 2014;2014:210241.
19. Svensson-Arvelund J, Mehta RB, Lindau R, Mirrasekhian E, Rodriguez-Martinez H, Berg G, et al. The human fetal placenta promotes tolerance against the semiallogeneic fetus by inducing regulatory T cells and homeostatic M2 macrophages. *J Immunol*. 2015;194(4):1534-1544.
20. Fiore S, Newell ML, Trabattoni D, et al. Antiretroviral therapy-associated modulation of Th1 and Th2 immune responses in HIV-infected pregnant women. *J Reprod Immunol*. 2006;70(1):143-150.
21. Fontas E, Van Leth F, Sabin C, et al. Lipid profiles in HIV-infected patients receiving combination antiretroviral therapy: are different antiretroviral drugs associated with different lipid profiles? *J Infect Dis*. 2004;189(6):1056-1074.
22. Squires J, Bricker DD, Twombly E. Ages & stages questionnaires: Paul H. Brookes Baltimore, MD, USA; 2009.
23. Beck EJ, Shields JM, Tanna G, Henning G, De Vega I, Andrews G, et al. Developing and implementing national health identifiers in resource limited countries: why, what, who, when and how? *Glob Health Action*. 2018;11(1):1440782.
24. Boulle A, Heekes A, Tiffin N, et al. Data centre profile: the provincial health data centre of the Western Cape Province, South Africa. *Int J Popul Data Sci*. 2019;4(2):1-11.
25. Malaba TR, Newell ML, Madlala HP, Perez A, Gray CM, Myer L. Methods of gestational age assessment influence the observed association between antiretroviral therapy exposure, preterm delivery, and small-for-gestational age infants: a prospective study in Cape Town, South Africa. *Ann Epidemiol*. 2018; 28(12):893-900.

26. Madlala HP, Malaba TR, Newell ML, Myer L. Elevated body mass index during pregnancy and gestational weight gain in HIV-infected and -uninfected women in Cape Town, South Africa: association with adverse birth outcomes.*Trop Med Int Health*. 2020; 25(6);702-713
27. Gray CM, Ross MC, Chanzu N, et al. Antiretroviral therapy (ART) during pregnancy is associated with increased placental inflammation and small for gestational age neonates: a prospective study. *9th IAS Conference on HIV Science*, Paris, 2017.
28. Zulu MZ, Chanzu N, Fick L, et al., editors. Differential distribution of M1 and M2 Macrophages in the decidua and chorionic villi of placentas from HIV-1 infected mothers on combination antiretroviral therapy. *9th IAS Conference on HIV Science*, Paris, 2017.