The cascade of care for HIV-exposed infants in rural South Africa

Thesis presented for the degree of

Doctor of Philosophy

(Field of Study: Medical Statistics and Epidemiology)

Elizabeth Alice Christine Chappell

Institute of Clinical Trials and Methodology, UCL

Declaration

I, Elizabeth Alice Christine Chappell, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: ………………………………………………………   Date: …………………………………….

Abstract

Human Immunodeficiency Virus (HIV) in infants is not straightforward to diagnose, those born with HIV have high mortality, and those starting treatment at older ages have poorer outcomes, making the optimisation of care for HIV-exposed infants extremely important. The cascade of care is a tool used to describe the pathway from testing of HIV-exposed infants to successful treatment of those diagnosed with HIV, and, when used at a population-level, can identify gaps in healthcare systems and monitor improvements over time.

This thesis used routinely collected data to assess the cascade for an estimated 17,570 HIV-exposed infants born between 2010 and 2016 in the Hlabisa health sub-district, KwaZulu-Natal, South Africa, a rural area where the antenatal HIV seroprevalence was 48% in 2016.

The key cascade stages considered were the proportion of HIV-exposed infants who received an HIV polymerase chain reaction (PCR) test, the proportion of infants diagnosed with HIV who initiated antiretroviral therapy (ART), and the proportion of infants on ART who achieved viral suppression. A deterministic and probabilistic data linkage algorithm was developed to link, at an individual-level, data from demographic surveillance, PCR test data from the National Health Laboratory Service, and data from TIER.net, the national Department of Health ART surveillance system.

Results suggest that by two years of age, 88% of HIV-exposed infants born in the sub-district had been tested, and of those diagnosed with HIV, 65% had initiated ART, of whom 53% were still on ART, of whom 68% were virally suppressed.

Limitations to each of the sources of data used were described, which may have led to errors in the linkage and estimates of the cascade. Improvements to data collection systems are required, both for the quality and continuity of clinical care as well as for surveillance and research purposes.

Impact statement

In this thesis, I use the concept of the cascade of care for infants exposed to Human Immunodeficiency Virus (HIV) to describe the milestones from testing to viral suppression that individuals exposed to and diagnosed with HIV need to pass through, comparing results from different routine data sources. I focus on infants born between 2010 and 2016 in a rural sub-district of KwaZulu-Natal, South Africa, where approximately half of pregnant women have HIV.

In 2015, South Africa became the first high HIV burden country to introduce routine polymerase chain reaction (PCR) testing of HIV-exposed infants at birth. The aim was to ensure timely identification of infants with HIV, and early initiation of treatment, as half of those not on treatment die by 2 years of age. My results suggest that less than a quarter of infants who receive a negative test result at birth have a follow-up test result as recommended, which may result in some infants who acquire HIV never being diagnosed. This level of repeat testing is much lower than findings from other studies, all from single hospitals in urban areas, and reasons for this poor coverage should be explored so that improvements can be made. I gave an oral presentation of these results to fellow researchers and multidisciplinary healthcare professionals at the 10th International Workshop on HIV Pediatrics.

I also found variation by clinic in adherence to national guidelines across all stages of the cascade. Repeat HIV testing varied by clinic, as did the frequency of viral load and CD4 monitoring among those on treatment. Identification of practises in clinics which contribute to better adherence may help healthcare workers from other clinics improve their own processes, and in turn improve overall healthcare across the sub-district.

My comparison of different data sources revealed significant limitations of such data in estimating the cascade. I show how there is incomplete reporting in many of the data sources used, including TIER.net, Road-to-health booklets and the District Health Information System. Further, I demonstrate the difficulty of linking each of the different datasets without a unique identifier. I suggest ways in which data collection and reporting through these systems could be improved, which is important both for clinical care as well as enabling a breadth of health research.

My findings in the context of other research from South Africa suggest that it is currently difficult to assess the success of the birth testing cascade of care introduced in 2015. Methodological improvements could greatly enhance the accuracy of estimates, and help South Africa properly evaluate its strategy. They will also help guide other countries who are considering the implementation of birth testing.

Acknowledgements

Table of contents

List of tables

List of figures

List of frequently used acronyms and abbreviations

3TC	Lamivudine
ABC	Abacavir
ACCDB	Africa Centre Clinical Database
ACDIS	Africa Centre Demographic Information System
AHRI	Africa Health Research Institute
AIDS	Acquired Immune Deficiency Syndrome
ANCHSS	National Antenatal Sentinel HIV & Syphilis Survey Report
aOR	Adjusted Odds Ratio
ART	Antiretroviral Therapy
ARV	Antiretrovirals
cART	Combination Antiretroviral Therapy
CDC	Centers for Disease Control and Prevention
CI	Confidence Interval
d4T	Stavudine
ddI	Didanosine
DHIS	District Health Information System
DNA	Deoxyribonucleic Acid
EFV	Efavirenz
EID	Early Infant Diagnosis
EMA	European Medicines Agency
EMTCT	Elimination of Mother-To-Child Transmission
FDA	US Food and Drug Administration
FTC	Emtricitabine
GEE	Generalized Estimating Equations
HIV	Human Immunodeficiency Virus
HR	Hazard Ratio
IMCI	Integrated Management of Childhood Ilness
IQR	Interquartile Range
LPV/r	Lopinavir/ritonavir
MONARCH	Management and Optimization of Nutrition, Antenatal, Reproductive, Child Health and HIV Care
MTCT	Mother-To-Child Transmission
NHLS	National Health Laboratory Service
NFV	Nelfinavir

NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NRTI	Nucleoside Reverse Transcriptase Inhibitor
NVP	Nevirapine
OR	Odds Ratio
PCR	Polymerase Chain Reaction
PEPFAR	President's Emergency Plan for AIDS Relief
PI	Protease Inhibitor
PIP	Population Intervention Platform
PMTCT	Prevention of Mother-to-Child Transmission
QIC	Quasi-likelihood under the Independence model Criterion
RNA	Ribonucleic acid
SAMPTCTE	South African Prevention of Mother to Child Transmission Evaluation
sdNVP	Single dose Nevirapine
SSA	Statistics South Africa
TDF	Tenofovir Disoproxil Fumarate
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNICEF	United Nations International Children's Emergency Fund
WHO	World Health Organization
ZDV	Zidovudine

Chapter 1. Introduction

1.1. Overview of HIV

Human Immunodeficiency Virus (HIV) was first identified in the early 1980s, following investigations into cases of a number of rare diseases usually only reported in individuals with severe immunosuppression, such as *Pneumcystis carinii* pneumonia (now known as *Pneumocystis jiroveci* pneumonia) and Kaposi Sarcoma, in previously healthy men who had sex with men in the United States [1]. Since then, the HIV epidemic has grown, with 37.9 million people estimated to be living with HIV in 2018 [2].

HIV is a retrovirus, with its genetic code stored as single stranded ribonucleic acid (RNA). There are two major types of the virus, HIV-1 and HIV-2. HIV-1 is the predominant and most widespread form, and is the focus of this thesis. Hereafter, the general term HIV will be used to refer specifically to HIV-1. Retroviruses need host cells in order to replicate, and the main host cells for HIV are CD4 lymphocyte cells, which are the part of the immune system responsible for coordinating the body's immune response to infection. The virus attaches itself to and enters CD4 cells, releasing its RNA. This genetic material converts to DNA (deoxyribonucleic acid) and integrates itself with the host's own DNA, so that as the CD4 cells replicate, new genetic material from the virus is produced. The newly formed HIV virions are subsequently released back into the blood stream, and go on to infect other CD4 cells. The invasion of the CD4 cells in this way causes damage, and ultimately leads to their destruction [3].

The virus is present in all body fluids, but is at higher concentrations in blood, semen, vaginal fluids and breast milk, and thus the most common routes of transmission are sexual intercourse, sharing or re-use of needles or syringes, and vertical transmission from mother-to-child.

The natural history of HIV in adults is shown in Figure 1.1. During primary infection, typically the first 3 to 6 months, there are high levels of circulating and replicating virus, resulting in a high rate of CD4 cell infection and destruction. During this time, viral reservoirs (groups of cells which are infected with the virus but are inactive and not replicating) are established [4], and individuals may experience symptoms including fever and fatigue [5]. Viral load (the term used to refer to number of copies of HIV per millilitre of blood, copies/mL) subsequently falls, and reaches a steady set point. CD4 count (which is measured in cells per cubic millimetre of blood, cells/mm^3) may increase again, although usually to a level lower than prior to infection [6].

Following this, individuals enter a period of asymptomatic infection, which can last for up to 10 years. During this phase, individuals experience a slow and steady depletion of CD4 cells over time, and the gradual weakening of their immune system, leaving them vulnerable to infections.

As this continues, individuals reach the advanced stage of HIV infection, termed progression to AIDS (Acquired Immune Deficiency Syndrome), with a number of opportunistic infections and illnesses, such as candidiasis, HIV-related encephalopathy, and recurrent pneumonia, considered to be AIDS-defining. Without treatment, the median survival for an adult whose CD4 count has fallen to <50 cells/mm^3 is 12 to 18 months [7, 8].

As well as an increased risk of infection, individuals with HIV are at increased risk of non-AIDS morbidity, for example cardiovascular disease, as a result of chronic immune activation [9].

Figure 1.1 - Natural history of HIV in adults

Removed due to copyright restrictions
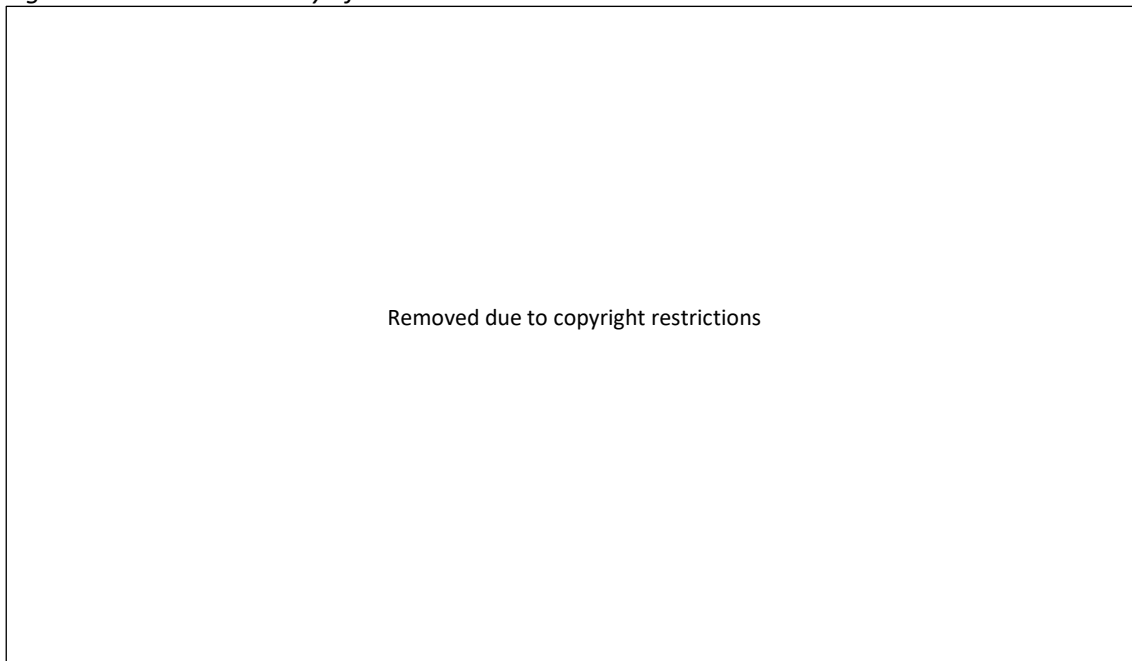
Reference: [10]

Antiretroviral therapy (ART) drugs (known as antiretrovirals, ARVs) can be used to prevent replication of the virus, reducing the amount circulating in the blood. The success of treatment is typically defined as the achievement of viral suppression below the limit of detection by a viral load assay. The detection limit varies according to the assay used, with more sensitive assays developed in more recent years, but thresholds are typically between <40 and <400 copies/mL. ARVs only stop HIV replicating and do not have any effect on the latent HIV reservoirs, so they are not able to cure individuals with HIV, however having an undetectable viral load limits further damage to the immune system, and individuals with a viral load <50 copies/mL are unable to transmit the virus to anybody else [11].

Different classes of drugs have been developed which prevent replication of the virus at different stages in its replication cycle. The first class developed was nucleoside reverse transcriptase inhibitors (NRTIs), with the first such drug (zidovudine, ZDV) licensed for use by the US Food and Drug Administration (FDA) in 1987 and others following shortly after [12].

Initially, these ARVs were given on their own (monotherapy) or later as two in combination (dual-therapy).

Drugs in two more classes, non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs), were subsequently developed and approved in the late 1990s. HIV has a fast replication cycle with a high genetic error rate, resulting in high genetic diversity, and under the selective pressure of ART this may result in the development of drug-resistant strains of the virus [13]. As a result, combination antiretroviral (cART) regimens, consisting of a backbone of two or more NRTIs plus either a PI or an NNRTI (often termed the third agent), became the standard of care. In more recent years, drugs from other classes have been developed, including entry inhibitors (EIs) and integrase inhibitors (IIs), although until the end of 2018 NNRTIs and PIs remained the preferred choice of third agent for first line regimens [14].

Initially, eligibility for ART was usually dependent on immunological and clinical criteria, with only patients who had more advanced disease progression receiving treatment. Over time, the thresholds for initiation were updated, widening access. In 2015, results from the START and TEMPRANO trials were published, demonstrating the benefit of immediate ART initiation even for individuals with high CD4 counts [15, 16], and the World Health Organization (WHO) subsequently recommended universal treatment for all individuals living with HIV [17].

Intellectual property rights give 20 year patents on newly developed pharmaceuticals [18], and the resulting market exclusivity means that new ARVs can be very expensive, limiting access to treatment in low- and middle-income countries. Further, high distribution costs and delays in licensing drugs in individual countries may restrict their availability. Initiatives by organisations such as the Medicines Patent Pool, which aims to broaden access to healthcare globally, led to an increased number of drug companies issuing voluntary licenses allowing the production and distribution of generic versions of their ARVs to low- and middle-income countries, even while still under patent. As a result, the price of first-line HIV treatment fell from $10,000 per person per year in 2000 [19] to less than $100 per year in 2015 [20], contributing to accelerated scale up of ART globally.

Virological failure on ART is typically defined as the repeated detection of the virus, and usually occurs because of poor adherence, but could also be the result of acquired resistance or drug-drug interactions. Although individuals experiencing such treatment failure may stay clinically and immunologically well, extended periods of high viremia while on treatment may lead to the development of new resistance, limiting future treatment options. Any adherence concerns should be addressed in the first instance, but individuals may eventually be switched to a second-line regimen.

Survival with HIV has improved dramatically since the start of the epidemic, with the life-expectancy of individuals diagnosed and initiating treatment in recent years approaching that of the general population [21].

1.2. HIV in infants and young children

1.2.1. Epidemiology of HIV in infants and young children globally

Globally, the total number of children under 14 years of age living with HIV was estimated to be 1.8 million in 2017, of whom 550,000 were aged under 5 [22]. The number of new infections among children aged under 14 years was estimated to be 170,000 in 2017, having fallen from 450,000 in 2000.

1.2.2. Mother-to-child transmission (MTCT)

HIV infection in children most commonly occurs through vertical transmission from mother-to-child, which can take place during pregnancy (in utero), during labour and delivery (intrapartum), or during breastfeeding (postpartum). In the absence of effective interventions it is estimated that the rate of MTCT is between 25-40% [23], with 10-20% of infections occurring during pregnancy, 50% during labour and delivery, and 30%-40% during breastfeeding [24, 25].

The strongest risk factor for MTCT is maternal viral load. An analysis from the Women and Infant Transmission Study (WITS) showed that the transmission rate increased from 1% among women with a viral load of ≤400 copies/mL at delivery to 32% among women with a viral load of >100,000 copies/mL [26]. Other risk factors include a low maternal CD4 count [27], younger gestational age at delivery [28], and complications during delivery such as the prolonged rupture of membranes [29].

Effective maternal ART during pregnancy and breastfeeding, as well as post-exposure prophylaxis for infants, are the most important interventions for the prevention of mother-to-child transmission (PMTCT). In 1994, results of the PACTG 076 trial demonstrated that giving ZDV orally to the mother from 14 weeks of pregnancy, intravenously during labour, and to the infant to 6 weeks of age reduced MTCT by 68% (with a transmission rate of 8.3% compared to 25.5% among those receiving a placebo) [30]. Subsequent trials and observational studies assessed alternative timing and durations of ZDV use, as well as the use of other drugs such as lamivudine (3TC) and nevirapine (NVP) [31], leading to widespread use of dual and triple therapy for PMTCT from the early 2000s.

In 2010, the WHO guidelines were updated with two possible options (A and B) for PMTCT for women (who were not otherwise eligible for ART for their own health) and their infants (Table 1.1) [32]. In order to simplify implementation in a setting with limited CD4 monitoring, in 2011

Malawi extended this and adopted their own Option B+, consisting of lifelong ART for the mother (regardless of their CD4 count) and daily NVP or ZDV for the infant for 4 to 6 weeks. This was recommended by the WHO in 2013 [33], and subsequently adopted by many countries, including low- and middle-income countries.

Table 1.1 - PMTCT guidelines for pregnant and breastfeeding women: WHO options A, B, B+

	Woman receives:		Infant receives:
	Treatment (CD4 ≤350 cells/mm^3)	Prophylaxis (CD4 >350 cells/mm^3)	
Option A	cART, starting as soon as diagnosed and continued for life	*Antepartum:* ZDV as early as 14 weeks gestation *Intrapartum:* sdNVP first dose of ZDV/3TC at onset of labour and *Postpartum:* Daily ZDV/3TC for 7 days	Daily NVP from birth to 1 week after the cessation of breastfeeding, or to 4-6 weeks if the mother is on treatment/not breastfeeding
Option B	cART, starting as soon as diagnosed and continued for life	cART early as 14 weeks gestation, continued to delivery/1 week after the cessation of breastfeeding	Daily NVP or ZDV from birth to 4-6 weeks of age
Option B+	Regardless of CD4 count, cART started immediately and continued for life		Daily NVP or ZDV from birth to 4-6 weeks of age

Adapted from [34]. sdNVP: single dose nevirapine

Other interventions for PMTCT include the primary prevention of HIV infection among young women, the prevention of unintended pregnancy among women with HIV, and regular antenatal testing for HIV, both at first presentation in antenatal care, throughout pregnancy and at delivery. Avoidance of breastfeeding is often recommended in high-income countries where safe alternative methods of infant feeding are available, but in resource-limited settings breastfeeding has important nutritional benefits and reduces the risk of diarrhoea and respiratory infections and subsequent mortality [35]. Transmission rates have been reported to be lower in infants who are exclusively breastfed compared to in those who are mixed fed [36]. Additionally, there is evidence that caesarean section prior to the onset of labour has a protective effect compared to vaginal delivery [37].

Global coverage of ART among pregnant women was estimated at 80% in 2017 [22], and the scale up of this and other PMTCT interventions has resulted in falling rates of transmission. Rates of less than 1% have been achieved in high-income countries [38], and of less than 5% reported in many low- and middle-income countries, including in UNAIDS (Joint United Nations Programme on HIV/AIDS) Global Plan priority countries (a set of 21 countries identified by UNAIDS in 2009 as accounting for 90% of the global number of pregnant women living with HIV) [39].

In 2014, the WHO defined criteria for the validation of the elimination of mother-to-child transmission (EMTCT) [40]. Targets consist of both impact criteria (≤50 new paediatric HIV infections per 100,000 live births and a transmission rate of either <5% in breastfeeding populations or <2% in non-breastfeeding populations) and process criteria (95% of pregnant women receiving antenatal care, 95% of pregnant women receiving HIV testing in pregnancy, and 95% of pregnant women diagnosed with HIV receiving treatment). In order for a country to receive validation of EMTCT, impact targets must be met for at least 1 year and process targets for at least 2 consecutive years. Cuba was the first country to receive validation in June 2015, followed by Thailand (the first country with a generalised HIV epidemic), Belarus and Armenia (2016), Anguilla, Antigua & Barbuda, Bermuda, Cayman Islands, Montserrat, and St. Christopher & Nevis (2017), Malaysia (2018) and the Maldives (2019) [41].

1.2.3. Diagnosis

Antibody assays are commonly used to detect HIV infection in adults and older children, but the diagnosis of HIV in infants is complicated by the presence of maternal HIV antibodies in all HIV-exposed infants usually until at least 12 months, but in some cases up to 18 months of age, regardless of whether or not they have actually acquired HIV [42]. Virological testing is therefore recommended to diagnose children under the age of 18 months [43], despite the associated increase in cost and the requirement for specialised equipment and trained technicians [44].

There are two types of virological assays for HIV, one for DNA and one for RNA. HIV RNA assays detect the amount of virus circulating in the blood, while HIV DNA is a marker of HIV persistence across different types of infected cells [45]. Although the sensitivity and specificity of RNA tests has been shown to be comparable to those of DNA tests for infants with no ART exposure [46], the use of ART in the mother or ART prophylaxis in the infant could result in a false-negative RNA test [47]. Only DNA PCR (polymerase chain reaction) tests are usually therefore recommended for diagnosis of infants, with quantitative RNA tests reserved for monitoring response to treatment among those with confirmed infection. Either whole blood specimens or dried blood spot samples can be used for DNA PCR testing. The sensitivity and specificity of DNA PCR tests by age, in infants receiving ART prophylaxis, are shown in Table 1.2.

Table 1.2 - Sensitivity and specificity of HIV DNA PCR tests in infants receiving ART prophylaxis, by age

	At birth	1 month	6 weeks	3 months
Sensitivity (probability of a positive result if infected)	55%	89%	98.8%	100%
Specificity (probability of a negative result if uninfected)	99.8%	100%	100%	100%

Results at birth, 1 month and 3 months of age taken from [46], and at 6 weeks of age from [48]

The WHO guidelines currently recommend routine testing of all HIV-exposed infants at 6 weeks of age [43]. There are however potential benefits to first testing infants at or soon after birth,

including access to a captive population of babies born within healthcare facilities and an opportunity to diagnose and initiate infants on ART prior to the peak of early mortality in the first 3 months of life [44]. Further, the use of more effective regimens and higher coverage of PMTCT means that proportionally more infections now occur in utero rather than intrapartum, resulting in a higher proportion of infants already infected with HIV being detectable at birth [49]. However, because not all infants with HIV would be detectable at birth, concerns about birth testing include the potential for missing some infants with HIV, so infants who test negative at birth therefore require an additional test after a few weeks to confirm their status. Due to the risk of false-positive results, the WHO recommends the receipt of a second confirmatory PCR test following an initial positive result. Additionally, a test at 6 weeks after the cessation of breastfeeding and an antibody test at 18 months of age may be recommended.

The timing of recommended PCR tests often corresponds to early vaccination and growth visits, which make convenient entry points for routine PMTCT testing. However, some infants, especially those born to mothers not previously diagnosed or those infected postpartum, may not be diagnosed until much later, often when they present sick at malnutrition units or hospitals [44].

Two point-of-care tests for infant diagnosis have been developed [50, 51] and early results suggest that they are accurate and lead to faster ART initiation [52, 53]. However, they are yet to be widely implemented across sub-Saharan Africa, and so testing is often reliant on central laboratories. The associated long turnaround time in getting samples to laboratories, for processing of the sample, and then getting the corresponding results back to clinics and caregivers creates an opportunity for the loss-to-follow-up of infants from care [54].

Global coverage of early infant diagnosis (EID), defined as the proportion of HIV-exposed infants who received a PCR test by two months of age, was estimated at 51% (95% CI, confidence interval, 41%, 67%) in 2017, an increase from 33% (27%, 45%) in 2010 [22].

1.2.4. ART

The biggest benefit of early treatment is the prevention of early mortality. Interim results from the CHER trial, which were published in 2008, demonstrated a 76% reduction in mortality after a median follow-up of 40 weeks for infants aged <12 weeks who initiated on ART immediately (among whom 4% died), compared to those who initiated when they met a clinical or immunological threshold (among whom 16% died) [55]. The results from the trial led to the WHO recommendation for immediate treatment initiation for all infants.

A second benefit of early treatment is the potential to prevent or reduce the latent viral reservoir that is formed during primary infection. This was evidenced by the case of the 'Mississippi baby'

(and other subsequent similar cases), who received triple ART within hours of birth resulting in early virological control followed by sustained virological remission off ART for several years (a so called functional cure), although virological rebound has since occurred [56]. Additional advantages to early treatment include the protection of long-term immunological reconstitution [57], better neurodevelopmental outcomes [58], and improved growth [59].

Despite these clear advantages, there are also some potential detriments of ART use in young children. With no cure available, ART use is currently expected to be life-long in vertically-infected children, leading to concerns over accumulating toxicity and resistance. In addition, treatment adherence is a concern long-term, as with many chronic conditions. There is therefore a growing interest in treatment sparing strategies, where early treatment is given, but following the achievement of viral suppression the number of drugs given is reduced [60]. Further studies have also demonstrated the potential benefit (at least in settings with regular virological monitoring) of complete interruptions to treatment [61, 62], and of other simplification strategies such as short-cycle (5 days on, 2 days off) therapy [63].

The paediatric drug development process is a barrier to successful ART treatment in very young children, with the development of ARV formulations suitable for infants and children lagging behind drug development for adults. Although incentives are offered, the paediatric approval process remains lengthy, and the relatively small (and shrinking) size of the population of children with HIV compared to of adults, especially in lucrative high-income country markets, means there is less motivation for drug development companies to invest in research into formulations for infants.

There are several issues with infant and paediatric drug administration which differentiate it from that in adults. Identification of appropriate doses for neonates, especially those born prematurely or at a low birth weight, is difficult as they have reduced hepatic and renal function, complicating drug metabolism [64]. Additionally, frequent recalculation of dosing is required as children grow, with failure to do this risking under dosing. WHO produces weight band dosing for many ARVs, though research usually starts with oldest and heaviest children and works down, meaning appropriate doses for infants are often the last to be identified. Paediatric formulations need to be easy to administer, with young children unable to swallow whole pills. In the past, the gold-standard paediatric formulation was thought to be oral liquids, with adjustments to dose made with the administration of more liquid. However, liquid formulations often require refrigerated storage, necessitating cold supply chains and appropriate places for families to store them, and accurately administering the correct amount can be difficult. Many liquids also have extremely poor palatability, making optimal adherence difficult. Bioavailability may change for some drugs when solid tablet formulations are manipulated, for example, area under the curve (AUC) decreases by 45% when solid lopinavir/ritonavir (LPV/r) tablets are

crushed [65], and so specialised formulations such as dispersible tablets or granules are therefore required.

As a result of these issues, there is limited availability of HIV drugs for young children, with only a quarter of ARVs approved by the European Medicines Agency (EMA) or the FDA for use in adults approved for children aged under 2 years [66].

A list of drugs approved for use in infants and young children is shown in Table 1.3. There are a number of issues with some drugs that are specific to their use in infants and young children. With regard to NRTIs, although abacavir (ABC) is recommended by the WHO for first-line treatment among children older than 28 days, there are limited safety data on its use in those under 3 months [67]. Tenofovir Disoproxil Fumarate (TDF) is licensed for use in children over 2 years, although it is not preferred in younger children, as it may adversely affect long-term bone and renal health [68]. The more recently developed Tenofovir Alafenamide may be safer in this respect, although data in children is limited [69]. Stavudine (d4T), didanosine (ddI) and nelfinavir (NFV) are no longer recommended due to safety and efficacy concerns. In 2002, the PENTA 5 trial demonstrated the superiority of a ABC+3TC backbone over ZDV+3TC [70], and this remains the preferred choice for younger children.

Table 1.3 - List of ARVs approved for use in infants and young children

Class	Drugs	Acronym	Weight/age restrictions (FDA licensing)
Nucelos(t)ide Reverse Transcriptase Inhibitors - NRTI	Abacavir	ABC	>3 months
	Didanosine*	ddI	
	Emtricitabine	FTC	
	Lamivudine	3TC	
	Stavudine*	d4T	
	Tenofovir Disoproxil Fumarate	TDF	>2 years
	Tenofovir Alafenamide	TAF	>25kg
	Zidovudine	ZDV	
Non-Nuceloside Reverse Transcriptase Inhibitors - NNRTI	Efavirenz	EFV	>3 months
	Etravirine	ETV	>2 years
	Nevirapine	NVP	
Protease Inhibitors - PI	Atazanavir	ATV	>3 months
	Darunavir	DRV	>3 years
	Lopinavir**	LPV/r	>14 days
	Ritonavir	RTV	
	Nelfinavir*	NFV	>2 years
	Fosamprenavir	FPV	
	Tipranavir	TPV	>2 years
Entry Inhibitors - EI	Maraviroc	MVC	>2 years
	Enfurvitide	T20	>6 years
Integrase Inhibitors - II	Dolutegravir	DTG	>20kg
	Elvitegravir	EVG	>25kg
	Raltegravir	RAL	>2kg
	Bictegravir	BIC	>25kg

*No longer recommended for paediatric use
**Boosted with low dose ritonavir

The choice between an NNRTI or PI depends on a number of factors. In general, PIs have a higher barrier to resistance, but usually require pharmacokinetic boosting with ritonavir and have more metabolic complications, while NNRTIs may confer cross-class resistance with a single mutation, but are cheaper than PIs and are less likely to interact with drugs taken for tuberculosis (TB) treatment. There is conflicting evidence from trials over which to use; PENPACT-1 reported similar outcomes for those on NNRTI and PI-based regimens [71], while in IMPAACT P1060, children on LPV/r did better than those receiving NVP [72].

There are specific issues with the use of some NNRTIs and PIs in younger children. LPV/r is not recommended for use in neonates less than 14 days old due to an increased risk of toxicities [73, 74]. Efavirenz (EFV) was not recommended in children under 3 years of age due to highly variable pharmacokinetics and poor absorption, although EFV sprinkles have recently been licensed by the FDA in infants over 3 months old on the basis of a pharmacokinetic model [68]. Due to concerns over transmitted drug resistance, the use of NVP (as well as other NNRTIs such as EFV) in infants who were previously exposed to NVP for PMTCT may not be recommended, especially in settings with no availability of baseline resistance testing; among children with NVP exposure, the IMPAACT P1060 showed higher treatment failure rates for those on a NNRTI compared to a PI [72].

1.2.5. Outcomes on ART

Infants and young children experience a different response to HIV infection compared to adults, with viral load peaking within 4 to 8 weeks and slowly declining over the first few years, compared to a lower peak and more rapid decline in adults [75]. Once initiated on ART, viral suppression in infants can take longer than in adults and older children, and they may experience higher rates of virological failure [76-78]. Concerns over switching too early and limiting future treatment options are particularly pertinent in young children, given their need for life-long treatment.

CD4 count varies substantially with age among children without HIV to at least 5 years of age, although CD4%, defined as the total CD4 cell count as a proportion of the total lymphocyte count, may be more stable [79]. Categorical stages are therefore often used to describe the immunological status of children with HIV, for example, the Centers for Disease Control and Prevention (CDC) 2014 immunological surveillance criteria, shown in Table 1.4 [80].

Without treatment, progression to AIDS in infants is quicker than in adults, with a more rapid CD4 decline likely a result of their immature immune system [81]. In addition, infants who acquire HIV in utero progress more quickly than those infected intrapartum [82], as do those born to women with more advanced disease [83]. Without treatment, one-third of infants living with HIV in sub-Saharan Africa die by their first birthday and half by their second birthday [62].

Among children on ART mortality may be high in the first year, driven by those initiating late with a low CD4 count, but among those on treatment long-term outcomes are generally good [84, 85].

Table 1.4 - CDC immunological stage based on CD4 count or CD4 percentage

Stage	<1 year		1-5 years		≥6 years	
	CD4 count, cells/mm^3	CD4%	CD4 count, cells/mm^3	CD4%	CD4 count, cells/mm^3	CD4%
1	≥1,500	≥34	≥1,000	≥30	≥500	≥26
2	750-1,499	26-33	500-999	22-29	200-499	14-25
3	<750	<26	<500	<22	<200	<14

CD4% defined as CD4 count as a proportion of total lymphocyte count

In addition to immunological staging, clinical staging may be used to describe disease progression among individuals with HIV; WHO clinical staging for children is described in Table 1.5 [86].

Table 1.5 - WHO clinical staging

WHO clinical stage	
1	Asymptomatic Persistent generalised lymphadenopathy
2	Unexplained persistent hepatosplenomegaly Recurrent or chronic upper respiratory tract infections (otitis media, otorrhoea, sinusitis, tonsillitis) Herpes zoster Lineal gingival erythema Recurrent oral ulceration Papular pruritic eruption Fungal nail infections Extensive wart virus infection Extensive molluscum contagiosum Unexplained persistent parotid enlargement
3	Unexplained moderate malnutrition not adequately responding to standard therapy Unexplained persistent diarrhoea (14 days or more) Unexplained persistent fever (above 37.5°C, intermittent or constant, for longer than one 1 month) Persistent oral candidiasis (after first 6 weeks of life) Oral hairy leukoplakia Lymph node tuberculosis Pulmonary tuberculosis Severe recurrent bacterial pneumonia Acute necrotizing ulcerative gingivitis or periodontitis Unexplained anaemia (<8g/dl), neutropenia (<0.5 x 10^9/l) or chronic thrombocytopenia (<50 x 10^9/l) Symptomatic lymphoid interstitial pneumonitis Chronic HIV-associated lung disease, including bronchiectasis
4	Unexplained severe wasting, stunting or severe malnutrition not responding to standard therapy Pneumocystis (jirovecii) pneumonia Recurrent severe bacterial infections (such as empyema, pyomyositis, bone or joint infection, meningitis, but excluding pneumonia) Chronic herpes simplex infection (orolabial or cutaneous of more than 1 month's duration or visceral at any site) Oesophageal candidiasis (or candidiasis of trachea, bronchi or lungs) Extrapulmonary tuberculosis Kaposi sarcoma Cytomegalovirus infection (retinitis or infection of other organs with onset at age more than 1 month) Central nervous system toxoplasmosis (after the neonatal period) HIV encephalopathy Extrapulmonary cryptococcosis, including meningitis Disseminated nontuberculous mycobacterial infection Progressive multifocal leukoencephalopathy Chronic cryptosporidiosis (with diarrhoea) Chronic isosporiasis Disseminated endemic mycosis (extrapulmonary histoplasmosis, coccidioidomycosis, penicilliosis) Cerebral or B-cell non-Hodgkin lymphoma HIV-associated nephropathy or cardiomyopathy

1.3. The cascade of care

The cascade of care is a model of the stages that an individual with HIV needs to pass through in order to achieve viral suppression, and is commonly used at a population-level to assess the performance of HIV healthcare systems. Intermediate stages are used to identify reasons why individuals fail to achieve viral suppression, and most commonly include diagnosis and initiation of ART, but linkage to care, eligibility for ART (where relevant), retention in care, and adherence to treatment, among others, may also be used. Achievement of the final stage of the cascade is important from both a public health perspective, to reduce the potential for onwards transmission [11, 87], and for the individual themselves, to allow immune recovery [15]. The concept of transition through a healthcare system was first used in 1972 in studies of hypertension care [88], and has been used since 2011 in HIV [89].

In 2014 UNAIDS announced the 90-90-90 treatment targets, with the aim that by 2020, 90% of people living with HIV should know their status, 90% of those diagnosed should have initiated ART, and 90% of those on ART should be virally suppressed [90]. In 2016, Sweden, a high-income and low HIV burden country, became the first to report meeting the 90-90-90 targets [91]. Several other countries and regions have reported to have met or be close to meeting the targets, including low- and middle-income countries with generalized HIV epidemics such as Botswana [2]. Results from mathematical models suggested that achievement of this target would end the AIDS epidemic by 2030 [90], however progress varies substantially, and at the end of 2017 the global cascade was only estimated to be 75-79-81 [22].

There are two broad methods for constructing cascades; cross-sectional and longitudinal. Cross sectional cascades, such as those designed to assess attainment of the 90-90-90 targets, summarise the proportion of people in a defined population who are at each stage of the cascade at a given point in time. Cascades constructed in this way are relatively easy to understand, and require fewer data and less complicated analyses [92]. Alternatively, a cascade constructed longitudinally, that is, by following a cohort of people over time, can provide additional insights, enabling the estimation of the proportion of individuals who have reached a stage a given number of years after infection or the median time to move between stages. From this it is possible to directly estimate how a patient would be expected to pass through the stages, without making an assumption of temporality within the population of interest (that is, that the probability of movement between stages has not changed over time) which would be required to estimate this with a cross-sectional cascade [92]. Further, longitudinal cascades are better suited to identifying predictors of moving (or failing to move) between stages.

Comparing the results of cascades from different settings can be complicated by inconsistent or sub-optimal methodology and definitions for each stage [93, 94]. Despite global and regional

attempts to standardise definitions [93, 95, 96], methodology still varies considerably. A recent systematic review compared 53 national cascades reporting data for 2010 to 2016, and of the 48 which specified the methodology used, nearly one-third were graded as having low-quality methodology (defined as being based on modelling studies or non-representative samples) [97].

There are a number of specific difficulties with the definition and measurement of each stage [93]. The total number of individuals living with HIV has to be estimated rather than measured, and this can be done using one of several different methods, for example, the use of population based surveys [98] or modelling [99]. All methods require numerous assumptions to be made and so commonly estimates for this stage are not available [100]. Various definitions of linkage to and retention in care may be used, depending on the data available. The definition of ART initiation can include all those ever initiated on ART, or be restricted to those with only long-term ART use, or those currently receiving ART, and further criteria may restrict the definition to just those receiving specific regimens [101]. Historically, there are challenges over how to account for those not yet eligible to start treatment, with some analyses choosing to add eligibility as an additional stage [102]. There is also variation in the definition of viral suppression. A Canadian cascade showed substantial differences in results when specific criteria regarding the threshold, duration and population eligible for the definition of viral suppression were changed [99]. Further, viral load monitoring is not always routinely done in low- and middle-income settings, with measurements often taken infrequently and targeted at those suspected of poor adherence or treatment failure, leading to questions of how to consider those with no measurement available. A review of recently published national cascades found 35% (28/81) were not able to estimate the proportion of individuals on ART who were suppressed [97]. Across the whole of the cascade, there is the additional challenge of how to consider individuals who have been lost to follow-up, have migrated, or who are known to have died.

As well as disparities in the definition of each cascade stage, estimates may vary depending on the source of data used. Routinely collected data are often cheaper and more convenient to use than collecting new data specifically to estimate the cascade, but their limitations may not be well considered. Disparate data sources are often used, meaning a different set of individuals are assessed at each stage of the cascade, or even that different unlinked datasets are used for estimates of the numerator and denominator of a given stage, which may impact the validity of conclusions [92].

Traditionally, HIV cascades have ended with the achievement of viral suppression, however subsequent stages such as immunological status [101] and health-related quality of life [103] may also be included, broadening the definition of achieving an optimal health state.

Despite these difficulties, the concept of the cascade of care is a useful tool to summarise and bring clarity to complex data [104], and has been widely used as the focus of targets to measure and improve the quality of HIV care.

1.4. South Africa

1.4.1. Demographics of South Africa

The population of South Africa was estimated at 51.7 million in 2011, of whom 29% were under the age of 15 years, and 11% were under the age of 5 years [105]. The country is divided into nine provinces (Figure 1.2), of which KwaZulu-Natal (where this thesis will focus) has the second largest population, estimated at 10.3 million in 2011.

Figure 1.2 - Map of South Africa
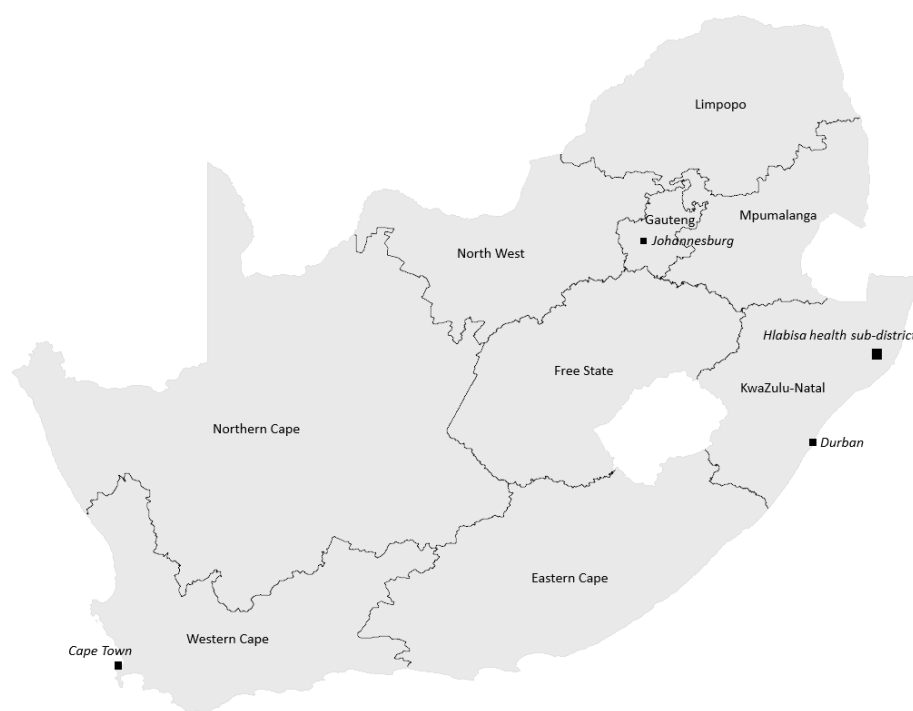
Although the World Bank classifies South Africa as an upper-middle income country, there is huge inequality, both by race and geographical area, and South Africa has the highest Gini coefficient (a measure of disparity in wealth distribution) of any country in the world [106]. There are large discrepancies in poverty, socio-economic status and access to basic healthcare by province, much of which is a legacy of the apartheid era [107, 108]. More than half the population lived below the national poverty line in 2015, a proportion that has been increasing over time. In 2017, unemployment was estimated to be 27%, and youth unemployment to be 39% [106].

Infant mortality in South Africa had fallen to 45.7 per 1,000 live births in 1993, but subsequently rose again as the HIV epidemic grew, reaching 53.6 in 2002, before subsequently declining again

to 30.0 in 2016. In 2007, infant mortality ranged by province from 25 per 1,000 live births (in the Western Cape) to 60 (in KwaZulu-Natal and the Eastern Cape) [108]. Life expectancy at birth was estimated at 62.3, 52.2, 57.1 years in 1990-95, 2005-10, 2010-2015 respectively, compared to 64.5, 68.8, 70.5 years globally [109]. Fertility rates (defined as the average number of children per woman) were 2.55 and 2.40 in 2005-2010 and 2010-2015 respectively, compared to 2.56 and 2.51 globally [109].

Across South Africa, 84% of infants born between 2014 and 2016 were ever breastfed (and was lower in KwaZulu-Natal at 71%), and the median duration of breastfeeding was 10.4 months (9.6 in KwaZulu-Natal) [110].

1.4.2. Epidemiology of HIV in South Africa

The first AIDS-related death in South Africa was reported in December 1981, after which the prevalence rose slowly over time to 0.8% in 1990 and 4.3% in 1994 [111]. Relative to other comparable countries, South Africa was slow to act on its HIV epidemic, in part because of resistance from senior political figures. Thabo Mbeki, who became President of South Africa in 1999, demonstrated strong opposition to HIV policy, disputing estimates of HIV prevalence in South Africa and denying the relationship between HIV and AIDS [112]. This resulted in delays to the implementation of a national policy for the rollout of ART, and contradictory public health messages around prevention and the need for treatment. Wide distribution of ART eventually began in 2004, but delays contributed to the loss of an estimated 2.65 million life-years [113].

South Africa has the largest HIV epidemic of any country globally, with an estimated 7.1 million people living with HIV in South Africa in 2016, equating to approximately 14% of the population [22]. Of this total, 320,000 individuals were aged between 0 and 14 years. There is high geographical variation in HIV prevalence by province, with the lowest prevalence seen in the Northern Cape (5%) and the highest in KwaZulu-Natal (17%) [114]. The number of new infections per year (among all ages) peaked in 1998 at 560,000, but has fallen over time to 270,000 in 2016 [22]. Among children aged less than 14 years, the estimated number of new infections peaked in 2003 at 70,000, and has decreased over time to 15,000 in 2016 (most dramatically between 2008 and 2010 from 56,000 to 26,000).

The number of AIDS-related deaths per year peaked in 2005 at 260,000 among all ages combined and 46,000 among children under 14 years of age, and has since fallen to 100,000 and 9,500 respectively in 2016. The number of children (aged <17 years) who were orphans because of HIV was estimated at 1.4 million in 2016.

South Africa now has the world's largest HIV treatment programme with an estimated 3.9 million people on ART, accounting for 20% of those on treatment globally [22]. ART coverage

increased from 0% in 2004 to 31% in 2011 [115]. In 2017, South Africa's progress towards the UNAIDS 90-90-90 targets was estimated to be 90-68-77 across the population as a whole, and 78-74-78 among children aged under 15 years [2]. Access to public primary healthcare facilities is free for all South African citizens, while access to secondary and tertiary care is free for some groups of patients (including pregnant women and those with HIV), with the level of charges for other patients dependent on their income [116].

1.4.3. South African national HIV guidelines

<u>PMTCT</u>

The national guidelines for PMTCT are shown in Table 1.6. There was initially reluctance from the South African government to introduce a PMTCT program, but following a court challenge by the Treatment Action Campaign, sdNVP was offered to mothers with HIV and their infants from 2002. For women not otherwise eligible for treatment for their own health, option A was introduced from 2010, and option B from 2013. Option B+ has been recommended since 2015, with coverage of ART among pregnant women estimated at 91% in 2014-2015 [114].

Table 1.6 - South African PMTCT guidelines over time

		Mother	Infant
2002 [117]		sdNVP in labour	sdNVP within 72 hours of birth
2004 [117]		ZDV from 28 weeks gestation, sdNVP in labour	sdNVP within 72 hours of birth
2008 [118]		ZDV from 28 weeks gestation, sdNVP in labour	sdNVP within 72 hours of birth, ZDV for 7 days or 28 days if mother received <4 weeks of ZDV
2010 [119]	Option A	ZDV from 14 weeks gestation, sdNVP and TDF/3TC in labour	NVP for 6 weeks if mother on ART for own health or if infant formula-fed, else NVP to end of breastfeeding
2013 [120]	Option B	ART for all pregnant women, stopped at cessation of breastfeeding if not otherwise eligible for own health	NVP for 6 weeks if mother on ART for own health or if infant formula-fed, else NVP to end of breastfeeding
2015 [121]	Option B+	Lifelong ART for all pregnant and breastfeeding women	NVP for 6 weeks, plus AZT if maternal viral load>1,000 copies/mL, or NVP for 12 weeks if mother received <4 weeks of ART by delivery

sdNVP: single dose nevirapine

<u>EID</u>

A summary of the South African EID guidelines is shown in Table 1.7. In line with the WHO, from 2004 guidelines recommended routine DNA PCR testing of all HIV-exposed infants at 6 weeks of age, with subsequent tests at 6 weeks after the cessation of breastfeeding and at any other time

for infants presenting symptomatically additionally recommended. The use of an HIV RNA viral load to confirm the status of any infants initially testing PCR positive was recommended.

In 2015, South Africa became the first high HIV-burden and low-/middle-income country to recommend routine testing of HIV-exposed infants at birth, with an immediate confirmatory PCR test for those testing positive (and immediate ART initiation, without waiting for the confirmatory result), and a repeat test at either 10 or 18 weeks of age (depending on the duration of NVP prophylaxis received) for those testing negative. Birth testing guidelines were introduced in April in KwaZulu-Natal, and June in all other provinces, apart from the Western Cape, which delayed implementation until April 2016 having previously already recommended targeted birth testing of high-risk infants [122].

Table 1.7 - South African EID guidelines over time

	2004	2008	2010	2013	2015
HIV-exposed and symptomatic	HIV PCR test at presentation				
HIV-exposed and asymptomatic	HIV PCR at ≥6 weeks	HIV PCR at 6 weeks			HIV PCR at birth
HIV PCR positive	HIV viral load at baseline, repeat HIV PCR only if asymptomatic		HIV status confirmed by viral load >10,000 copies/ml	HIV status confirmed by any quantifiable viral load	Confirmatory HIV PCR
HIV PCR negative	Repeat HIV PCR test if infant symptomatic, and at 6 weeks after cessation of breastfeeding				
					Repeat at 10 weeks (or 18 weeks if received 12 weeks NVP) or if breastfed and maternal viral load >1,000 copies/ml

Adapted from [123].

Treatment guidelines

Comprehensive rollout of ART in South Africa began in 2004, although demonstration projects providing ART for patients with advanced disease had been running in a small number of clinics since 2001 [124]. Higher priority for treatment was initially given to those who had lower CD4 counts, but access to treatment was increased with revisions to the guidelines in 2010, 2013 and 2015, and in September 2016 the South African government announced the introduction of universal treatment in line with WHO recommendations (Table 1.8). Treatment for all infants

under 1 year of age has been recommended since 2010, and for all children <5 years of age since 2013.

Table 1.8 - South African HIV treatment guidelines over time: when to start

	2004 [125]	2010 [126]	2013 [120]	2015 [121]	2016 [127]
<1 year	CD4<20%; WHO stage 2/3; recurrent/ prolonged hospitalisations	All	All	All	All
1-5 years	CD4<15%; WHO stage 2/3; recurrent/ prolonged hospitalisations	CD4≤750/25%; WHO stage 3/4	All	All	All
5-15 years	CD4<15%; WHO stage 2/3; recurrent/ prolonged hospitalisations	CD4≤350; WHO stage 3/4	CD4≤350; WHO stage 3/4	CD4≤500; WHO stage 3/4	All
≥15 years	CD4≤200; WHO stage 4	CD4≤200; CD4≤350 and TB, pregnant* or WHO stage 4; MDR-/XDR-TB	CD4≤350; WHO stage 3/4; TB; pregnant/ breastfeeding*; meningitis	CD4≤500; WHO stage 3/4; TB; pregnant/ breastfeeding**; HBV	All

HBV: Hepatitis B virus; MDR-TB: Multi-drug resistant tuberculosis; XDR-TB: Extensively drug-resistant tuberculosis
Individuals meeting any one of the criteria separated by a semicolon (;) were eligible for ART
WHO stage refers to WHO clinical staging
*For duration of pregnancy/while breastfeeding only (if not otherwise eligible)
**Then continued life-long

A summary of first-line regimens recommended for infants and young children in the national guidelines is shown in Table 1.9. Since 2010, ABC+3TC+LPV/r has been recommended for those under 3 years of age and ABC+3TC+EFV for those over 3 years (with the option to replace EFV for LPV/r for those with NVP exposure for PMTCT, because of concerns around NNRTI resistance). No specific recommendations are given for neonates under 1 month of age, with guidelines suggesting healthcare providers seek specialist supervision.

Since 2010, guidelines have recommended switching infants and young children failing NVP- or EFV-based regimens to a LPV/r-based one, concurrent with a change to the NRTI backbone. Expert advice is recommended when choosing second-line regimens for those failing on LPV/r. Access to third-line regimens is restricted, being controlled by the National Department of Health.

Table 1.9 - South African HIV treatment guidelines over time: first line regimens for infants and children

	2004 [125]	2010 [126]	2013 [120]	2015 [121]
<28 days	Specialist supervision required, no recommendations given*	Specialist supervision required, no recommendations given	Specialist supervision required, no recommendations given	Specialist supervision required, no recommendations given
28 days - <3 years (or <10kg)	d4T+3TC+LPV/r (or can swap LPV/r for NVP if not used in PMTCT)	ABC+3TC+LPV/r	ABC+3TC+LPV/r	ABC+3TC+LPV/r
>3 years and >10 kg	d4T+3TC+EFV/NVP	ABC+3TC+EFV	ABC+3TC+EFV (or can swap EFV for LPV/r if NVP exposed in PMTCT for >6 weeks)	ABC+3TC+EFV (or can swap EFV for LPV/r if NVP exposed in PMTCT for >6 weeks)
Notes		Keep on d4T if no side-effects	Swap anyone still on d4T to ABC	Swap anyone still on d4T/ddI to ABC

*Applied to initiation of ART in all infants up to 6 months of age

Viral load/CD4 monitoring

A summary of the national guidelines for viral load and CD4 monitoring for children and infants on ART is shown in Table 1.10. Since 2010, viral load monitoring has been recommended at 6 and 12 months after ART initiation. Subsequent testing is then recommended every 12 months while suppressed (every 6 months between 2010 and 2015 for those aged <5 years), and every 2 months (every 3 months until 2013) for those with viremia. CD4 monitoring is recommend at 6 (until 2013) and 12 months after initiation, then every 12 months thereafter if clinically indicated.

Infant vaccination and feeding

A summary of the recommended vaccination schedule for all infants (regardless of HIV exposure or infection status) in South Africa since April 2009 is shown in Table 1.11. Vaccinations are recommended at birth, 6, 10 and 14 weeks, 9 and 18 months, and 5 years of age [128]. The South African National Department of Health routinely provided free formula feed for HIV-exposed infants until August 2011, and recommended all mothers to breastfeed their infants until 12 months of age from April 2013 [120, 129].

Table 1.10 - South African viral load/CD4 monitoring guidelines over time

	2004 [125]	2010 [126]	2013 [120]	2015 [121]
Viral load monitoring				
<5 years of age	Every 6 months	At 6 and 12 months after ART initiation. Then every 6 months while suppressed <1,000 copies/mL, else every 3 months	At 6 and 12 months after ART initiation. Then every 6 months while suppressed <1,000 copies/mL, else every 2 months	At 6 and 12 months after ART initiation. Then every 12 months while suppressed <1,000 copies/mL, else every 2 months
>5 years of age		At 6 and 12 months after ART initiation. Then every 12 months while suppressed <1,000 copies/mL, else every 3 months	At 6 and 12 months after ART initiation. Then every 12 months while suppressed <1,000 copies/mL, else every 2 months	
CD4 monitoring				
All ages	Every 6 months	At 6 and 12 months after ART initiation. Then every 12 months if clinically indicated.	At 12 months after ART initiation. Then every 12 months if clinically indicated.	At 12 months after ART initiation. Then every 12 months if clinically indicated.

Table 1.11 - South African childhood vaccination schedule to 5 years of age from April 2009

Age of child	Vaccination
Birth	BCG Polio
6 weeks	Polio DTP + HIB Hepatitis B Rotavirus PCV (pneumococcal)
10 weeks	Polio DTP + HIB Hepatitis B
14 weeks	Polio DTP + HIB Hepatitis B Rotavirus PCV (pneumococcal)
9 months	Measles PCV (pneumococcal)
18 months	Polio DTP Measles
5 years	Polio DT

DTP+HIB: Diphtheria, tetanus, acellular pertussis, and haemophilus influenza type B; DT: Diphtheria and tetanus

1.5. Hlabisa health sub-district

The Hlabisa health sub-district is located in northern KwaZulu-Natal (Figure 1.2), within the uMkhanyakude district, which is the second most deprived district in South Africa. The sub-district is predominately rural, but contains an urban township with several peri-urban settlements. There is high unemployment, with 67% of those aged 18 years of age and over unemployed in 2010 [130], and high population mobility, with 22% of residents reporting in 2011 to have moved at least once in the last two years, of which 73% of migration events were to a destination outside of the sub-district [131]. The most recent estimate of the population, made in 2011, was 228,000 [132], and population density varies from 20-3,000 people/km^2 with an average household size of 7.9 members [133].

There are 16 primary public healthcare clinics in the sub-district, along with one public hospital. The hospital has 275 beds, and services offered include emergency care, medical, surgery, maternity, paediatrics, and HIV and TB clinics. Among the 10 clinics with data available, the median (interquartile range, IQR) [range] number of nurses working at the clinic is 8 (6, 11) [6, 3], and number of days a doctor is scheduled to attend the clinic each month is 1 (1, 1) [1, 4].

The majority of HIV patients attend clinic every month to collect prescriptions of ART, though recently attempts have been made to reduce clinic burden by introducing medicine pick-up points for stable patients at other locations, for example in pharmacies [134]. The number of ART visits per month at each clinic varies considerably (ranging from 200 to 4,500 per month in 2016), but across all constitute approximately one-third of the total number of clinic visits for any health condition. ART delivery is nurse-led, and there are limited paediatric- or adolescent-specific services. The overall median travel time to a clinic was 81 minutes in 2006 (although this is likely to have reduced with the creation of several new clinics since this analysis), with nearly 40% of individuals reporting the use of some form of public transport for at least part of the route, and clinic usage has been observed to decline with increasing travel time [135].

In 2016, HIV prevalence in the sub-district was estimated to be 37% [136]. Among men, the incidence rate was relatively stable until 2012, when it declined from 2.49 per 100 person-years to 1.01 (95% CI 0.58, 1.76) in 2017 [137]. Among women, it was stable to 2014, after which it declined from 4.89 (4.09, 5.84) to 3.06 (2.38, 3.94) in 2017. Incidence was higher among women aged 15-29 years (compared to older women) with less of a decrease over time, although prevalence was higher among women aged 30-49 years, reaching 59% in 2017. The antenatal seroprevalence was estimated at 47.5% between July 2015 and January 2017 [138]. By 2014, life expectancy had increased by 15.2 years to 45.9 years for men, and by 17.2 years to 54.2 years for women since the roll-out of ART in 2004 [139]. A study looking at the cascade of care for individuals with HIV aged ≥15 years found that 8 years after first testing positive only 45% had

been linked to care, 35% (of the total denominator) had initiated ART and 33% (of the total denominator) had therapeutically responded (defined as either viral suppression or achievement of good immunological status) [102].

There are limited data available on children with HIV within the sub-district, especially in recent years. The child (defined as <5 years of age) and infant (defined as <1 year of age) mortality rates (including those both with and without HIV) within the sub-district have decreased significantly over time, reaching approximately 20 and 38 per 1,000 live births respectively in 2014 [140, 141]. Approximately one-third of deaths in those under 5 were HIV-related, the majority of which occurred at home after seeking care from a healthcare provider [142]. The most recently published data on ART use among children followed 477 individuals who had started treatment aged ≤15 years to the end of June 2008 [143]. At ART initiation, the median age was 6.2 years, and 67% were WHO clinical stage III or IV (the most severe). By the end of follow up, 32 (7%) had died and 18 (4%) were lost to follow up. A study to the end of 2007 used deterministic modelling to estimate that only two-thirds of children requiring ART had initiated treatment [144]. In 2008, a program of training in paediatric HIV for clinic staff led to an increase in the number of children and infants starting ART [145]. Among a sample (n=101) of paediatric patients failing ART in 2011-2012, 91% were found to have a drug resistant mutation [146].

Three large studies in pregnant women have taken place within the Hlabisa health sub-district. The Hlabisa pregnancy cohort assessed the PMTCT cascade among pregnant women in care between 2010 and 2014; 43% of women with HIV were already on or initiated life-long ART within 6 months of their first antenatal visit, with coverage improving over time [147]. At 6 weeks post-partum, 80% of women were still in follow-up. Incident HIV during pregnancy was estimated at 4.5 new infections per 100 person-years, which, after adjustment for age and other demographic factors, was lower than in non-pregnant women [148]. The Vertical Transmission Study followed 2,722 pregnant HIV-negative and HIV-positive women living in the Hlabisa health sub-district from 2001 to 2005. The primary outcome of the study was to determine the effect of infant feeding practices on transmission rates, and the overall post-natal HIV transmission risk (between 4-6 weeks and 6 months) was found to be 4% among women who exclusively breastfed [149, 150]. The MONARCH trial (conducted between 2015 and 2017) assessed the effect of a quality improvement intervention on PMTCT process, which was found to have a positive impact on maternal viral load screening but not repeat HIV testing [138].

1.6. Objectives and overview of thesis

Children born with HIV have high mortality and those starting treatment at an older age have poorer outcomes, so it is important to identify infants who have acquired HIV early and link them to care, to enable the timely initiation of ART and achievement of viral suppression. The cascade

of care is a useful tool for describing the milestones individuals have to pass through in order to achieve this viral suppression, and at a population-level can be used to evaluate the performance of a healthcare system. The antenatal seroprevalence across South Africa is high, approaching 50% in some areas, including in the Hlabisa health sub-district, so although the rate of MTCT has fallen over time, many infants are still born with HIV (or acquire it later through breastfeeding), making the optimisation of care for HIV-exposed infants in this setting extremely important. There is no routine monitoring of EID systems in South Africa, and so high quality research is required to understand failures in the current system, enabling them to be rectified.

The overall objective of this thesis is to assess the cascade of care for infants exposed to HIV in the Hlabisa health sub-district, South Africa. I will focus on outcomes up to 2 years of age, because this covers the use of DNA PCR testing and by this time most infants would be expected to have been diagnosed and linked to care. However, for the later stages of the cascade in particular, outcomes over a longer duration of follow-up will also be reported. Although the cascade of care has traditionally begun with the denominator of individuals with HIV, here its use has been broadened to cover all infants exposed to HIV, starting from the testing of HIV-exposed infants to successful treatment of those diagnosed with HIV.

More specifically, the objectives of this thesis are to:

- Undertake data linkage to assemble the datasets required to measure the stages of the cascade in infants exposed to HIV
- Estimate the cascade of care for HIV-exposed infants born within the Hlabisa sub-district, with particular focus on the following stages:
  - Estimate the proportion of HIV-exposed infants who receive a PCR test, and the proportion of infants infected with HIV who are diagnosed
  - Among infants diagnosed with HIV, estimate the proportion who go on to initiate ART
  - Among infants initiating ART, estimate the proportion achieving viral suppression
- Compare management of infants exposed to and diagnosed with HIV across the cascade with recommendations made in the South African national guidelines
- Discuss the advantages and disadvantages of the use of routinely collected data in estimating the cascade

The population of interest is all HIV-exposed infants born in the Hlabisa health sub-district to women with HIV between June 2010, when immediate ART for all infants was introduced in South Africa, and December 2016.

This thesis has ten chapters. In Chapter 2, I review the literature around the cascade of care for HIV-exposed infants in South Africa. In Chapter 3, I describe the main sources of data used throughout the thesis, and in Chapter 4, I describe the algorithm I developed to link these datasets to enable the construction of a cascade of care. In Chapter 5, I compare four different methods and data sources in estimating the proportion of HIV-exposed infants who ever received a PCR test, and in Chapter 6, I use two of these to investigate factors associated with the receipt of PCR testing. In Chapter 7, I describe the frequency, timing and results of PCR testing in the sub-district among those who ever received a test, with particular reference to trends over time before and after the introduction of birth testing in South Africa in 2015. In Chapter 8, I investigate the proportion of infants diagnosed with HIV who went on to initiate ART, and their subsequent outcomes on ART including viral suppression, treatment interruptions and switches, and mortality. In Chapter 9, I bring together the results from the previous chapters to estimate the cascade of care. Finally, a summary and discussion of the key findings is given in Chapter 10.

Parts of the results from Chapter 7 were presented as an oral presentation at the 10th International Workshop on HIV Pediatrics [151], and parts of the results from Chapter 8 were presented as a poster at the 11th International Workshop on HIV Pediatrics [152].

Chapter 2. Literature Review

2.1. Introduction

In this chapter, I summarise the literature to date investigating the cascade of care for HIV-exposed infants in South Africa. I focus on the three most commonly used stages of the cascade, being the proportion of HIV-exposed infants who receive HIV PCR testing, the proportion of infants diagnosed with HIV who go on to initiate ART, and the proportion of infants on ART who achieve viral suppression.

2.2. Methods

South Africa is unlike many other high HIV burden countries, being classified by the World Bank as upper-middle income. In addition, it is the only high burden country to recommend PCR testing for infants at birth, rather than in the first few weeks of life. As a result, only studies based in South Africa (or across multiple regions including South Africa, with results disaggregated by country) were included in the review.

A separate literature review was conducted for each of the following three topics:

1. the proportion of HIV-exposed infants receiving a PCR test
2. the initiation of ART among infants diagnosed with HIV
3. the achievement of viral suppression among infants initiating ART

In order to ensure the studies summarised were relevant to my thesis, only those carried out over a time period that reflected the national guidelines from 2010 onwards were included. For topic 1, studies conducted prior to 2004, when routine PCR testing was introduced, were excluded. Until 2010 (or 2008 in the Western Cape), ART initiation of infants aged under 1 year was dependent on clinical and immunological criteria, therefore studies prior to this were excluded from topic 2, as ART coverage would be expected to be lower, and topic 3, to ensure comparable populations of infants on ART were included.

Studies had to report outcomes among infants up to two years of age (including those also covering older children, but with results disaggregated by age). Only studies reported in English, and those reporting quantitative research were included. The literature review included published studies to the 28th September 2019.

The primary method of searching for literature was using PubMed, a database of biomedical literature. The search terms used for each of the three topics are given in Table 2.1, with the same search terms used for topics 2 and 3.

Table 2.1 - Search terms used to identify studies in PubMed

Topic	Search terms
1 PCR testing	(HIV[tiab] OR AIDS[tiab] OR "human immunodeficiency virus"[tiab]) AND (infant*[tiab] OR neonat*[tiab]) AND ("South Africa"[tiab]) AND ("HIV Infections/diagnosis"[majr] OR "early infant"[tiab] OR "early diagnosis"[tiab] OR "EID"[tiab] OR "birth testing"[tiab] OR PCR[tiab] OR "polymerase chain reaction"[tiab] OR "testing"[tiab] OR "prevention of mother-to-child transmission"[tiab] OR PMTCT[tiab] OR "prevention of vertical transmission"[tiab])
2 ART initiation 3 Viral suppression	(HIV[tiab] OR AIDS[tiab] OR "human immunodeficiency virus"[tiab]) AND (infant*[tiab] OR neonat*[tiab] OR child*[tiab]) AND ("South Africa"[tiab]) AND (ART[tiab] OR antiretroviral[tiab] OR outcomes[tiab])

[tiab] field tag instructs PubMed to search the title and abstract of each record

In addition to searching PubMed, databases or abstract booklets (where applicable) from the last 3 years of the following conferences were searched, using the search term "South Africa" to identify potentially relevant literature: the International Workshop on HIV Pediatrics (2017, 2018 and 2019), the International AIDS Conference (2018 only, as the conference was not held in 2017 or 2019), the International AIDS Society Conference on HIV Science (2017 only, as the conference was not held in 2018 and the abstract book from 2019 was not available online at the time this review was conducted), and the Conference on Retroviruses and Opportunistic Infections (2017, 2018 and 2019). Finally, studies identified for each topic were cross-checked for the inclusion of outcomes relevant to the other topics, and their reference lists were checked for additional studies not otherwise identified.

2.3. PCR testing

2.3.1. Results

Figure 2.1 describes the studies identified for the review of PCR testing coverage (topic 1). Using the search terms described in the methods, 272 papers were identified through PubMed. Of these, 199 were excluded following review of their title and abstract, with the most common reason for exclusion being not estimating PCR testing coverage (n=117). The remaining 73 papers were fully reviewed, resulting in the exclusion of 52, with the main reason again being not reporting estimates of testing coverage (n=35). Three papers were excluded as they estimated testing coverage at the same location and using the same methods as three other studies, but covered a shorter time period [153-155]. One paper was excluded as it covered the time period before the introduction of routine PCR testing in South Africa, and instead looked

at the coverage of 12 month antibody testing [156]. A further three studies were found following searches of conference databases and two from a review of other papers, giving a total of 26 studies to be included. One conference abstract that covered the same analysis as one of the papers but contained extra analyses was also included, but the two were counted as a single study in this review [157, 158].

Figure 2.1 - Flowchart of studies identified for the review of HIV PCR testing coverage (topic 1)

```
                    ┌─────────────────────────┐
                    │ 272 papers identified   │
                    │ from PubMed             │
                    └─────────────────────────┘
                              │
                              │──────────────────────────────────────────────┐
                              │                                               │
                              │   Excluded following review of title/abstract: 199
                              │   No coverage estimates: 117
                              │   Not HIV: 34
                              │   Review article: 16
                              │   Qualitative: 13
                              │   Study protocol: 8
                              │   Modelling: 6
                              │   Case studies: 2
                              │   Not South Africa: 2
                              │   Cohort profile: 1
                              │
                    ┌─────────────────────────┐
                    │ 73 papers reviewed      │
                    └─────────────────────────┘
                              │
                              │   Excluded following review of paper: 52
                              │   No coverage estimates: 35
                              │   Modelling: 9
                              │   Duplicate: 3
                              │   Results not disaggregated by country: 2
                              │   Review article: 2
                              │   Before introduction of PCR testing: 1
                              │
                    ┌─────────────────────────┐
                    │ 21 papers included      │
                    └─────────────────────────┘
   ┌──────────────────────────────────┐
   │ Additional studies identified: 5 │
   │ Conferences abstract: 3          │
   │ Review of other papers: 2        │
   └──────────────────────────────────┘
                              │
                    ┌─────────────────────────┐
                    │ 26 studies included in  │
                    │ review                  │
                    └─────────────────────────┘
```

A summary of the outcomes and characteristics of the 26 studies included in the review is shown in Table 2.2. Three studies reported the proportion of infants who ever received a PCR test, 8 reported the proportion tested at birth or receiving a follow-up test following a negative result at birth, 14 reported the proportion tested by or at 6 weeks of age (excluding those looking specifically at testing at birth, but including one which estimated coverage using two methods, hence the numbers for this column summing to 15), and 4 reported the proportion receiving a repeat test following a positive or indeterminate result. Across all four outcomes, the most

common source of data was patient records or National Health Laboratory Service data (NHLS, the service responsible for all laboratory testing in public healthcare facilities in South Africa), and, excluding studies looking at a national or provincial level, most studies were based in urban areas.

Table 2.2 - Characteristics of the 26 studies included in the review of HIV PCR testing coverage (topic 1)

	Proportion of HIV-exposed infants:			
	Ever receiving a PCR test	Receiving a PCR test at birth, and a follow-up PCR testing following a negative result at birth	Receiving a PCR test at/by 6 weeks of age	Receiving a confirmatory PCR test following a positive result or a repeat PCR test following an indeterminate result
	Table 2.3	Table 2.4	Table 2.5	Table 2.6
Total number of studies	3	8	14**	4
Primary source of data				
Patient records	1	5	5	1
NHLS	2	2	3	3
NHLS plus road-to-health booklet	0	1	0	0
Interventional study	0	0	3	0
Interview with caregiver	0	0	2	0
Road-to-health card/booklet	0	0	1	0
DHIS	0	0	1	0
Geographical coverage				
Single healthcare facility	0	3	5	1
Group of healthcare facilities	1	3	6	0
Provincial	2	1	0	1
National	0	1	4	2
Setting*				
Urban	1	6	6	1
Rural	0	0	5	0
N/A	2	2	4	2

*Relevant to studies at a single healthcare facility or group of healthcare facilities only
**One study estimated coverage with two different methods, and so numbers in this column sum to 15
DHIS: District Health Information System; NHLS: National Health Laboratory Service

Proportion of HIV-exposed infants who ever received a PCR test

Three studies looked at the proportion of HIV-exposed infants who ever received a PCR test, with their study design and results described in Table 2.3. Two, by Smith et al (2019) and Moodley et al (2013), were conducted by the same research group, and looked at testing coverage for whole of KwaZulu-Natal over different time periods (the first from 2005 to 2011, and the second from 2013 to 2015) [157-159]. The studies used the same method, multiplying estimates of the number of live births (from birth registration data collected by Statistics South Africa, SSA) and of the seroprevalence (from a national survey, National Antenatal Sentinel HIV & Syphilis Survey Report, ANCHSS) in the province to give the denominator of the number of HIV-exposed infants, and using NHLS data to estimate the number of infants who received a PCR

test. Across the two studies, the estimated coverage increased from 15% in 2005 to over 100% in 2015. Although a laboratory barcode was used to deduplicate infants with repeat tests when estimating the number of infants tested for the numerator, the authors concluded that the overestimation in recent years was likely caused by poor linkage of repeat tests. The third study, by Kalk et al (2018), estimated the proportion of infants ever tested among those born at a group of healthcare facilities in Cape Town between 2013 and 2016 to be 89%, based on data recorded in patients' notes [160].

Proportion of HIV-exposed infants who received a PCR test at birth, and received a follow-up test after a negative result at birth

Eight studies looked at the proportion of HIV-exposed infants who received a routine PCR test at birth (that is, excluding those looking at targeted birth testing), or at the proportion of infants with a negative result at birth (either routine or targeted) who subsequently received a follow-up test (Table 2.4) [122, 160-166].

Four studies looked at the proportion of HIV-exposed infants receiving routine PCR testing at birth [122, 160, 162, 163]. The first three studies all looked at the coverage of laboratory-based (as opposed to point-of-care) birth testing [122, 160, 163]. Two studies, by Kalk et al (2018) and Technau et al (2017), which were both based in urban areas, each looked at coverage during a pilot program of birth testing (prior to the introduction of birth testing into the national guidelines) using data from patient records [160, 163]. In the study by Technau et al (2017), which included all HIV-exposed infants born at a hospital in Johannesburg, coverage was high at 98% [163]. In the other, by Kalk et al (2018), which was based at a group of primary care clinics in the Western Cape, only 61% of HIV-exposed infants were tested at birth [160]. A third study, by Moyo et al (2018), estimated the coverage of birth testing at a national level between June 2015 and May 2016 using NHLS data. They calculated the number of tests conducted in infants younger than 7 days of age divided by an estimate of number of exposed infants born, although they were not able to deduplicate repeat tests in the numerator [122]. Coverage was 76% over the whole time period, increasing from 39% in June 2015 (a month after the introduction of birth testing into the national guidelines) to 93% in May 2016. Across the whole time period, coverage in KwaZulu-Natal, where birth testing was introduced earlier (April 2015), was higher at 92%. The fourth study, by Moyo et al (2019), looked at the coverage of point-of-care PCR tests [162]. During weekday working hours (8am-4pm), point-of-care tests were offered to 3,475 infants born in 4 tertiary obstetric units. Coverage varied by site from 46% to 81%, however it was unclear whether infants received routine laboratory PCR testing outside of these hours.

Table 2.3 - Estimates of the proportion of HIV-exposed infants ever receiving a PCR test

Study (First author, publication year)	Time period covered	Setting	Data sources	Study design	Number of HIV-exposed infants	Number and/or proportion of infants ever receiving a PCR test
NHLS data						
Smith, 2019 [157, 158]	Jan 2013- Dec 2015	KwaZulu-Natal	NHLS, SSA, ANCHSS	NHLS PCR tests deduplicated using linkage based on unique identifiers. Number of HIV-exposed infants estimated from antenatal seroprevalence (ANCHSS) multiplied by number of live births (SSA).	Range: 94,836 to 98,250 per year	92% in 2013, 90% in 2014, >100% in 2015
Moodley, 2013 [159]	Jan 2005- Dec 2011	KwaZulu-Natal	NHLS, DHIS, ANCHSS	NHLS PCR tests deduplicated using linkage based on unique identifiers. Number of HIV-exposed infants estimated from antenatal seroprevalence (ANCHSS) multiplied by number of live births (DHIS).	Range: 13,699 to 70,284 per year	Increased from 18% in 2005 to 97% in 2011
Patient records						
Kalk, 2018 [160]	May 2013- Jun 2016	Primary care facility and referral centres, Cape Town	Patient records	All infants born to women with HIV included, including some during periods of targeted birth testing	2,012	1,794 (89%)

ANCHSS: National Antenatal Sentinel HIV and Syphilis Report; DHIS: District Health Information System; NHLS: National Health Laboratory Service; SSA: Statistics South Africa

Six studies looked at the proportion of infants testing negative at birth who ever received a follow-up test [122, 160, 161, 164-166]. The first three studies used data from patient records [160, 164, 165]. In the previously mentioned primary care clinic in the Western Cape by Kalk et al (2018), although birth testing coverage was relatively low, the majority (80%) of those tested returned for a subsequent test [160]. The coverage of follow-up testing was similar among infants who received targeted birth testing in same clinics during an earlier time period (79%) [160]. Comparable estimates for subsequent testing were observed among two other studies in infants undergoing birth testing, both also in urban areas [164, 165]. Nelson et al (2017) reported that 78% of infants who received a laboratory-based birth test and 75% of those who received a point-of-care birth test at primary care clinic in Khayelitsha were re-tested [164]. In a second study by Dunning et al (2017), 73% of infants who received targeted birth testing at a single hospital in Cape Town returned for subsequent testing [165].

Repeat testing in the final three studies was estimated using PCR test data from the NHLS database [122, 161, 166]. The first, by Maritz et al (2016), was conducted across the Western Cape during the time that targeted birth testing was recommended there, and found that 49% of infants were re-tested [161]. The second, by Moyo et al (2018), was a national study conducted following the introduction of routine birth testing, which reported an increase in coverage from 13% in June 2015 to 48% in May 2016 [122]. Estimates from this study would likely be an overestimate of repeat testing however, as they were based on the total number of infants receiving a test at 10 weeks of age and not restricted to those who previously received a test at birth. The final study, by Mazanderani et al (2018), was a pilot study conducted across Tshwane district, Gauteng [166]. A unique identifier was assigned to infants with their road-to-health booklet (which is issued to all infants born in South Africa and includes health-related data from throughout childhood), and healthcare workers were instructed to enter this number onto laboratory test request forms. Of 5,309 infants with a birth PCR test with a road-to-health booklet identifier in the NHLS data, 635 (12%) were found to have another PCR test, as indicated by the same identifier recorded for a later test. Coverage here was likely to have been underestimated because of healthcare workers failing to enter the road-to-health booklet identifier on all test request forms, with one available on only 39% of all birth tests conducted.

Table 2.4 - Estimates of the proportion of HIV-exposed infants receiving a PCR test at birth, and of follow-up PCR testing following a negative result at birth

Study (First author, publication year)	Time period covered	Setting	Data sources	Study design	Number of HIV-exposed infants	Number and/or proportion of HIV-exposed infants receiving birth testing	Number and/or proportion with a follow-up test after negative result at birth	Median (IQR) time to follow-up test
NHLS								
Moyo, 2018 [122]	Jun 2015-May 2016	National	NHLS	Birth coverage estimated as the number of tests conducted <7d divided by the estimated number of HIV-exposed infants born. Repeat testing estimated as proportion of HIV-exposed infants tested at 10 (±2) weeks of age	263,202	76% nationally (increasing from 39% to 93% by month), 92% in KwaZulu-Natal	Increased from 13% to 48% by month	Not reported
Maritz, 2016 [161]	Jan 2009-Jun 2015	Western Cape	NHLS	Targeted birth tests conducted <7d probabilistically linked to subsequent tests	3,322	N/A	1,547 (49%)	Mean 52 (SD 0.6)
NHLS plus road to health booklet identifier								
Mazanderani, 2018 [166]	May 2016-May 2017	Tshwane district, Gauteng	NHLS	Unique patient identifier distributed with road-to-health booklet entered into NHLS	5,309	Not reported	625 (12%)	Not reported

…Table 2.4 continued

Patient records								
Moyo, 2019 [162]	June- Dec 2018	Four tertiary obstetric units, Gauteng	Patient records	Point-of-care tests offered during set hours	3,475	46%-81% by site	Not reported	Not reported
Technau, 2017 [163]	Jun 2014- Dec 2016	Single hospital, Johannesburg	Patient records	Pilot birth testing program, all HIV-exposed infants eligible	6,861	6,695 (98%)	Not reported	Not reported
Kalk, 2018 [160]	May 2013- Nov 2015	Primary care facility and referral centres, Cape Town	Patient records	All infants born to women with HIV included, received targeted birth testing	456	N/A	358 (79%)	6.4w (6.0, 7.1)
	Dec 2015- Jun 2016			All infants born to women with HIV included, received routine birth testing	349	211 (61%)	168 (80%)	8.6w (6.4, 10.7)
Nelson, 2017 [164]	Nov 2014- Jul 2015	Primary care clinic, Khayelitsha	Patient records	All HIV-exposed infants who received birth testing processed through a central laboratory	436	Not reported	336 (78%)	44d (42, 47)
	Aug 2015- Apr 2016			All HIV-exposed infants who received point-of-care birth testing	321	Not reported	241 (75%)	45d (43, 50)
Dunning, 2017 [165]	Jul 2013- Aug 2015	Single hospital, Cape Town	Patient records	All infants born to women with HIV included, received targeted birth testing	551	N/A	401 (73%)	60d

IQR: Interquartile range; NHLS: National Health Laboratory Service; SD: Standard Deviation

Proportion of HIV-exposed infants receiving a PCR test at/by 6 weeks of age

Fourteen studies assessed the proportion of infants who received a PCR test by or at around 6 weeks of age, not including those looking at the proportion tested at birth, with estimates ranging from 0.4% to over 100% (Table 2.5) [123, 167-179].

Three studies looked at coverage at a national level [123, 167, 168], with one study making two estimates based on different data sources [123]. Of these, two (by Sherman et al (2014) and Sherman et al (2017)) estimated the numerator as the number of PCR tests conducted in infants under 2 months of age in the NHLS database, divided by an estimate of the number of exposed infants born made using estimates of the antenatal seroprevalence multiplied by the number of live births [123, 168]. The proportion of infants tested in the first study increased from 0.4% in 2003 to 73% in 2012 [168], and in the second increased from 52% in 2010 to 87% in 2014 [123]. The authors were unable to deduplicate multiple PCR tests conducted on the same infants, which would have led to the overestimation of coverage. Sherman et al (2017) also estimated the proportion tested using data from monthly clinic reports to the Department of Health (through the District Health Information System, DHIS), with a higher proportion reported, increasing from 87% in 2010 to 103% in 2014 [123]. The third study, by Diallo et al (2016), estimated coverage as the ratio of tests in NHLS conducted by 6 weeks of age compared to the total number of tests conducted, that is, they did not count HIV-exposed infants who never received a PCR test in the denominator [167]. The estimated coverage increased slightly over time, from 57% in 2011 to 62% in 2015.

In two studies, interviews were conducted with caregivers to determine each infant's testing status [169, 170]. Horwood et al (2010) conducted a study at 27 rural clinics in KwaZulu-Natal, where 164 women with HIV were asked immediately following immunization clinic visits whether their infant had received a PCR test, with 78 (48%) reporting that they had [170]. Reported coverage was compared to that recorded in their infant's Road-to-health booklet, where the test was only recorded in 65 (40%) cases. Feucht et al (2014) conducted a retrospective study, with caregivers of 201 children aged ≤7 years seen at an urban hospital asked whether their child had received a PCR test at 6 weeks of age [169]. Only 52% of caregivers reported the receipt of a test, with the results from this study likely affected by recall bias, which may have led to either under or overestimation.

Three studies assessed coverage before and after the implementation of an intervention, with each using data from patient records in clinics [171-173]. Two studies were conducted in rural areas and one in an urban hospital, and all reported higher coverage with the intervention. Schwartz et al (2015) compared infants PCR testing coverage at 6 weeks of age at a primary healthcare clinic in Johannesburg between those whose mother received a mobile phone based

intervention, among whom 76% were tested, to a pre-intervention cohort control group, among whom 45% were tested [171]. Doherty et al (2009) implemented a PMTCT quality improvement intervention in clinics, which led to an increase in the proportion of HIV-exposed infants tested at 6 weeks of age from 24% to 68% [172]. Tomlinson et al (2014) carried out a cluster-randomized trial, with 74% of infants visited by community health workers receiving a PCR test, compared to 67% of those in the control group [173]. As it was a cluster randomized trial, the interpretation of the difference between the two groups would not be confounded by changes in testing over time, unlike the studies by Doherty et al (2009) and Tomlinson et al (2014), which both used a historical control group. Given the interventional nature of all three studies, their generalisability to routine clinical practice may be limited.

Nsibande et al (2013) used data from infants' road-to-health cards (used prior to Road-to-health booklets) collected for a sub-study of a trial. Of a relatively small sample size of 48 HIV-exposed infants whose card was reviewed, only 3 (6%) had a PCR test recorded at their 6 week postnatal visit [174].

The final five studies were based on data recorded in patient records [175-179]. Two, by Geddes et al (2008) and Chetty et al (2012), looked at the proportion tested at the same hospital in Durban [178, 179]. The first estimated coverage between 2004 and 2007 to be 82% [179], with similar coverage observed between 2008 and 2009 in the second (83%) [178]. The third study, by Lilian et al (2013) was conducted at a single hospital in Johannesburg, where, as well as patient records, the NHLS database was searched for tests conducted at other sites [177]. In this study, 82% of HIV-exposed infants born between 2008 and 2010 were tested. A fourth study, by Fatti et al (2014), looked at coverage across three urban hospitals in the Eastern Cape between 2009 and 2012, and reported that 51% of HIV-exposed infants were tested [176]. The final study by Smith et al (2014) was the only one not conducted in an urban area, and used data from a single rural hospital and its feeder clinics in KwaZulu-Natal, and reported 54% of infants born in 2012 received a PCR test [175].

Proportion of HIV-exposed infants receiving a confirmatory PCR test following a positive result and a repeat PCR test following an indeterminate result

Four studies looked at confirmatory testing following an initial positive PCR result (Table 2.6) [157, 158, 163, 180]. In a study by Technau et al (2017) at a hospital in Johannesburg, 96/99 (97%) of infants identified as having HIV through a pilot birth testing program received a subsequent test [163]. The other two studies each used a linkage algorithm to identify repeat tests within the NHLS database, one across the whole of South Africa and one across KwaZulu-Natal. In the national study, by Mazanderani et al (2018), the proportion re-tested varied between 42% and 55% between January 2010 and December 2015 with no trend over time,

although the median time to confirmatory test decreased from 241 days to 17 days [180]. In the KwaZulu-Natal study, by Smith et al (2019), coverage of confirmatory testing increased from 16% in January 2013 to 31% in December 2015 [157, 158].

Only one study, by Mazanderani et al (2017), looked at repeat testing following an indeterminate result (Table 2.6), and used data linkage to identify repeats tests within the NHLS database. In this analysis, 34% of infants had a repeat test, at a median 29 (interquartile range, IQR 13, 57) days later [181].

2.3.2. Summary

The proportion of HIV-exposed infants ever receiving a test increased over time, however in recent years estimates were very close to or exceeded 100%. Estimates based on data from NHLS were generally higher than from other sources; this may have been the result of not being able to deduplicate repeat PCR tests on the same infant (either at all or reliably) or because of unreliable data collection for estimation of the denominator. There was high variation in the estimates of the coverage of both 6 week and birth testing, although for both outcomes estimates from most studies were much lower than estimates of the proportion ever tested, at between 50% and 80% across both outcomes.

Coverage of repeat testing after a negative birth test was high (>73%) in three studies conducted using patient records from urban healthcare facilities, though much lower (reaching only approximately 50% by the end of the time period studied) in a national study based on NHLS data. Similarly, coverage of confirmatory testing after both positive and indeterminate results was low when estimated using NHLS data, but close to 100% at a single hospital in Johannesburg.

Table 2.5 - Estimates of the proportion of HIV-exposed infants receiving a PCR test at/by 6 weeks of age

Study (First author, publication year)	Time period covered	Setting	Data sources	Study design	Number of HIV-exposed infants	Number and/or proportion receiving a PCR test at/by 6 weeks of age
NHLS						
Diallo, 2016 [167]	Jan 2011-Dec 2015	National	NHLS	Coverage estimated as number of PCR tests conducted ≤6 weeks divided by the total number conducted.	Not reported	Increased from 57% in 2011 to 62% in 2015
Sherman, 2017 [123]	Jan 2010-Dec 2014	National	NHLS, SSA, ANCHSS	Numerator estimated as the number of PCR tests conducted ≤2 months according to NHLS (not deduplicated). Number of HIV-exposed infants estimated from antenatal seroprevalence (ANCHSS) multiplied by number of live births (SSA).	Approx. 270,000 per year*	Increased from 52% in 2010 to 87% in 2014
Sherman, 2014 [168]	Jan 2003-Dec 2012	National	NHLS, SSA, ANCHSS	Numerator estimated as the number of PCR tests conducted ≤2 months according to NHLS (not deduplicated). Number of HIV-exposed infants estimated from antenatal seroprevalence (ANCHSS) multiplied by number of live births (SSA).	Range: 240,000 to 270,000 per year	Increased from 0.4% in 2003 to 73% in 2012
DHIS						
Sherman, 2017 [123]	Jan 2010-Dec 2014	National	DHIS	Numerator is 'infant first PCR test conducted around 6 weeks'. Denominator is 'number of live births to HIV-positive women'.	Approx. 250,000 per year*	Increased from 87% in 2010 to 103% in 2014
Interview						
Feucht, 2014 [169]	Jun 2009-May 2010	Single hospital, Gauteng	Interview with caregiver	Retrospective study of all HIV-exposed children aged under 7 years in care at hospital	201	104 (52%)
Horwood, 2010 [170]	Oct 2007-Feb 2008	27 primary health clinics in rural KwaZulu-Natal	Interview with mother consultation following	All infants attending 6 week immunization clinic with mother who reported themselves to have HIV	164	78 (48%)

Intervention

Schwartz, 2015 [171]	May-Jul 2013	Single primary health clinic, Johannesburg	Patient records	Assess impact of a mobile phone based intervention on PCR testing coverage at 6 weeks. Compared coverage with intervention to a pre-intervention cohort.	Without intervention	50	22 (45%)
					With intervention	50	38 (76%)
Doherty, 2009 [172]	Jan 2007-Sep 2008	3 hospitals and 18 clinics in Amajuba district, rural KwaZulu-Natal	Patient records	Assess testing coverage before and after introduction of a quality improvement intervention package	Before	Not reported	Not reported (24%)
					After	Not reported	Not reported (68%)
Tomlinson, 2014 [173]	Jun 2008-Dec 2010	Umlazi, rural KwaZulu-Natal	Patient records	Cluster randomized trial assessing impact of community health worker home visits	Control arm	698	465 (67%)
					Intervention arm	571	420 (74%)

Road to health card/booklet

Nsibande, 2013 [174]	Jan 2008-Dec 2011	Umlazi, rural KwaZulu-Natal	Road-to-health cards	Road-to-health card data collected as a sub-study of a community based cluster randomised trial		48	3 (6%)

…Table 2.5 continued

Patient records						
Smith, 2014 [175]	Jan- Dec 2012	Single hospital and 8 feeder clinics, rural KwaZulu-Natal	Patient records	All HIV-exposed infants born at hospital and surrounding clinics included. Data collected to May 2013.	843	455 (54%)
Fatti, 2014 [176]	Jan 2009-Mar 2012	3 urban hospitals, Eastern Cape	Patient records	All HIV-exposed infants born at hospitals included	910	468 (51%)
Lilian, 2013 [177]	Aug 2008-Jul 2010	Single hospital, Johannesburg	Patient records, and searching NHLS for tests from other sites	All HIV-exposed infants born at hospital included	838	691 (82%)
Chetty, 2012 [178]	May 2008-May 2009	Single hospital, Durban	Patient records	All HIV-exposed infants born or receiving care at hospital included	264	220 (83%)
Geddes, 2008 [179]	Mar 2004-Feb 2007	Single hospital, Durban	Patient records	All HIV-exposed infants born at hospital included	699	571 (82%)

ANCHSS: National Antenatal Sentinel HIV and Syphilis Report; DHIS: District Health Information System; NHLS: National Health Laboratory Service; SSA: Statistics South Africa
*Figures not reported, approximated from graph

Table 2.6 - Estimates of the proportion of HIV-exposed infants receiving a confirmatory PCR test following a positive result and a repeat PCR test following an indeterminate result

Study (First author, publication year)	Time period covered	Setting	Data sources	Study design	Number of infants	Number and/or proportion with a repeat test	Median (IQR) time to repeat test
Confirmatory test following initial positive result							
Patient records							
Technau, 2017 [163]	Jun 2014-Dec 2016	Single hospital, Johannesburg	Patient records	Infants with HIV identified through a pilot birth testing program	99	96 (97%)	Not reported
NHLS							
Smith, 2019 [157, 158]	Jan 2013-Dec 2015	KwaZulu-Natal	NHLS	Repeat PCR tests identified using linkage based on unique identifiers	Range: 1,599 to 2,144 per year	Increased from 16% in 2013 to 31% in 2015	Not reported
Mazanderani, 2018 [180]	Jan 2010-Dec 2015	National	NHLS	Repeat PCR tests identified using linkage based on unique identifiers	Range: 10,310 to 15,538 per year	Range: 42% to 55% (no trend over time) per year	Decreased from 241 in 2010 to 17 days in 2015
Repeat test after indeterminate result							
NHLS							
Mazanderani, 2017 [181]	Jan 2013-Dec 2015	National	NHLS	Indeterminate results identified in NHLS database, which was then searched for subsequent tests using automated linking algorithm	49,654	16,895 (34%)	29d (IQR 13, 57)

NHLS: National Health Laboratory Service

2.4. ART initiation

2.4.1. Results

A description of the studies found for the review of ART initiation is shown in Figure 2.2. Using the search terms described in the methods, 781 papers were identified in PubMed. After reviewing their titles and abstracts, 717 were excluded, with the most common reasons being not estimating ART coverage (n=334) and not including infants (n=169). The full text of the remaining 64 was reviewed, following which an additional 56 were excluded, with the most common reason again being not estimating ART coverage (n=49). One paper estimated ART coverage at a hospital at Johannesburg during a time period which was covered by two other papers from the same hospital, and so was excluded [155]. A study from the review of testing coverage (topic 1) which also looked at ART coverage, but which was not identified here, was also included [165]. In total, 9 studies were included in the review of ART coverage. One study looked at ART coverage from 2005-2010 across the Western Cape; only data from 2008 onwards were included, as this in when immediate ART initiation for infants was recommended in the province [182]. Another study conducted between 2008 and 2010 at a single hospital in Johannesburg was also included as all diagnosed infants were immediately initiated on ART, despite this not being national guidelines at the time [177].

A summary of the characteristics of the 9 studies included in the review is shown in Table 2.7. The most common source of data was patient records (n=7), with one using data from NHLS and one from caregiver interviews. One study looked at coverage of ART across a whole province. The remaining 8 looked at a single healthcare facility or group of facilities, of which 6 were based in an urban area, 1 in a rural area, and 1 used data from a mixture of the two.

Table 2.7 - Characteristics of the 9 studies included in the review of ART initiation (topic 2)

	Studies reporting the proportion of diagnosed infants initiating ART
Total number of studies	9
Primary source of data	
Patient records	7
NHLS	1
Interview with caregiver	1
Geographical coverage	
Single healthcare facility	4
Group of healthcare facilities	4
Provincial	1
Setting*	
Urban	6
Rural	1
Mixture of urban and rural	1
N/A	1

*Relevant to studies at a single healthcare facility or group of healthcare facilities only

Figure 2.2 - Flowchart of studies identified for the review of ART initiation (topic 2)
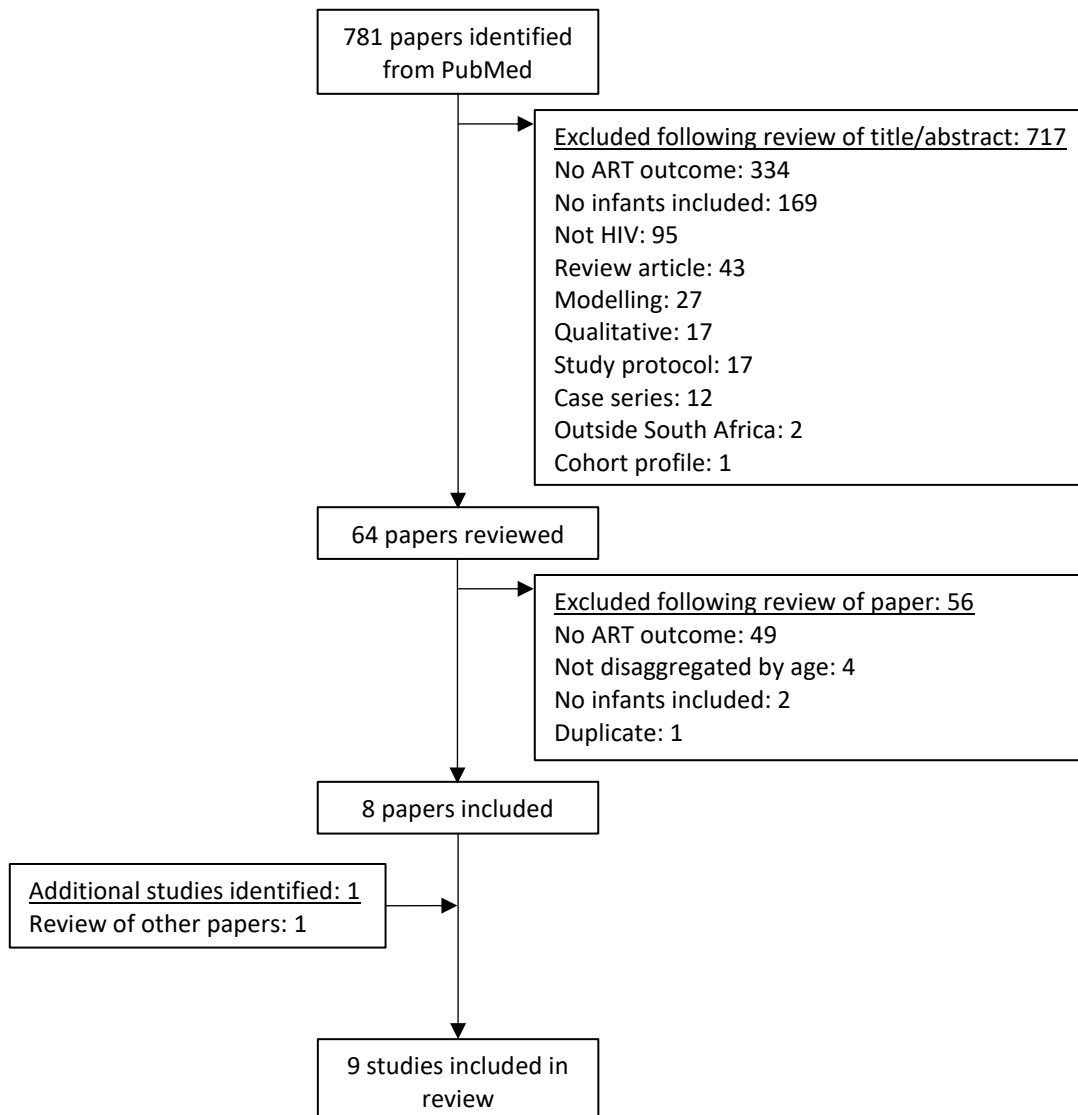
A summary of the design and results from the 9 studies included is given in Table 2.8. All of the studies reported estimates of the coverage of ART, and three reported the time to initiation from either the date of first positive PCR test or the date of the return of infant HIV test results to caregivers.

Table 2.8 - Estimates of the proportion of infants diagnosed with HIV who initiated ART

Study (First author, publication year)	Time period covered	Setting	Data source	Study design	Method of diagnosis	Number of infants diagnosed with HIV	Number and/or proportion initiating ART	Time to initiation
Patient records								
Technau, 2017 [163]	Jun 2014-Dec 2016	Single hospital, Johannesburg	Patient records	All infants diagnosed with HIV at hospital	Universal birth testing	99	95 (96%)	Not reported
Teasdale, 2019 [183]	Apr 2013-Dec 2015	5 urban healthcare facilities, Eastern Cape	Patient records	All infants diagnosed with HIV. Assess proportion initiating within 7 days.	Various	84	42 (50%)	Not reported
Dunning, 2017 [165]	Jul 2013-Aug 2015	Single hospital, Cape Town	Patient records	Infants with HIV identified through targeted birth testing matched to those diagnosed at 6 weeks. ART uptake by Feb 2016 assessed.	Birth test	21	18 (86%)	Not reported
					6 week test	4	4 (100%)	Not reported
Technau, 2018 [184]	Sep 2013-Jun 2014	Single hospital, Johannesburg	Patient records	All infants diagnosed with HIV at hospital	Targeted birth testing	13	13 (100%)	Not reported
Smith, 2014 [175]	Jan- Dec 2012	Single hospital and 8 feeder clinics, rural KwaZulu-Natal	Patient records	All infants diagnosed with HIV at healthcare facilities included	Majority at 6 week test	36	20 (56%)	Not reported
Abrams, 2017 [185]	Jan- Dec 2011	3 hospitals and 2 clinics, Johannesburg	Patient records	All infants diagnosed with HIV at healthcare facilities included	Various, including both inpatient and outpatient services	272	226 (83%)	From PCR test: 36 days (95% CI 31, 40) From receipt of test results: 23 days (95% CI 18, 27)
Lilian, 2013 [177]	Aug 2008-Jul 2010	Single hospital, Johannesburg	Patient records	All infants born and diagnosed at hospital included. Follow up to December 2010.	6 week test	38	23 (61%)	From PCR test: 10.1 weeks (IQR 1.9, 24.3) From receipt of test results: 6.1 weeks (IQR -3.3, 20.3)

Interview								
Mathivha, 2019 [186]	Oct 2012-Sep 2014	580 primary health clinics across South Africa	Telephone interview with caregiver	Infants diagnosed with HIV among a sample of HIV-exposed infants from each clinic.	PCR tests conducted at 6 weeks, then 3 monthly intervals to 18 months	59	2 (4%)	Not reported
NHLS								
Hsiao, 2013 [182]	Jan 2008-Dec 2010	Western Cape	NHLS data	Positive PCR tests in infants linked to viral load tests in those ≤5 years (recommended at ART initiation in Western Cape)	Majority at 6 week test	2,925	69% in 2008, 69% in 2009, 71% in 2010	From PCR test: Approx. 40 days*

NHLS: National Health Laboratory Service

*Figures not reported, approximated from graph

Seven studies estimated ART coverage using data recorded in patient records [163, 165, 175, 177, 183-185]. Only one study, by Smith et al (2014), looked at linkage to ART among infants in a rural area [175]. Among 36 infants diagnosed at a hospital (or its feeder clinics) in 2012, 56% went on to initiate ART and another 17% commenced ART work up but never actually started treatment. The other six studies were all based in urban areas [163, 165, 177, 183-185]. One study, by Teasdale et al (2019), was based at healthcare facilities across the Eastern Cape [183]. The study only reported the proportion of infants who had initiated ART within a week of diagnosis, at which time 50% were on treatment. The estimates from the other five studies ranged from 61% to 100% [163, 165, 177, 184, 185]. Three of the studies, by Lilian et al (2013), Technau et al (2018) and Technau et al (2017), were all conducted at the same hospital in Johannesburg [163, 177, 184]. Reported coverage increased from 61% among those diagnosed at 6 weeks of age between August 2008 and July 2010 [177], to 100% among those diagnosed through targeted birth testing between September 2013 and June 2014 [184] and 96% among those diagnosed through universal birth testing between June 2014 and December 2016 [163]. The fourth study, by Dunning et al (2017), was based at a single hospital in Cape Town and reported 86% (18/21) of infants diagnosed with a test at birth and 100% (4/4) of those diagnosed at 6 weeks of age went on to initiate ART [165]. The final study was conducted by Abrams et al (2017) [185]. Among 272 infants diagnosed in both inpatient and outpatient services at 5 healthcare facilities in Johannesburg during 2011, 226 (83%) initiated treatment. This high proportion initiating treatment may be a result of the sicker inpatient population included in the study.

Mathivha et al (2019) conducted a study across 580 primary healthcare clinics between October 2012 and September 2014, with caregivers of infants identified as having HIV contacted by telephone at least 3 months after diagnosis [186]. Although 101 infants with HIV were identified, the caregivers of only 59 could be contacted, at which time only 2 (4%) reported their infant to be on ART. The low coverage may be a result of selection or reporting bias, although these may have been expected to lead to higher estimates of ART use. The authors were unable to trace ART use in clinic registers for those not contacted.

Hsiao et al (2013) used NHLS data from across the whole of the Western Cape province, linking positive PCR tests to viral load tests (which were recommended at treatment initiation in the Western Cape at the time), with coverage steady at between 69% and 71% between 2008 and 2010 [182]. Coverage may have been underestimated because of healthcare workers failing to conduct the recommended viral load tests, or due to limitations of the data used for linkage.

Three of the studies reported time from diagnosis to initiation [177, 182, 185]. Among infants initiated on ART as part of the study based at a single hospital in Johannesburg between August 2008 and July 2010, the median (IQR) time from PCR test to initiation was 10.1 (1.9, 24.3) weeks

and from receipt of results to initiation was 6.1 (-3.3, 20.3) weeks, with the negative lower quartile suggesting some infants were presumptively started on treatment before diagnosis had been formally made [177]. In the study by Hsiao et al (2013) based on NHLS data from across the Western Cape, the median (IQR) time between PCR test and first viral load (a proxy for ART initiation) was approximately 40 days [182]. Finally, in the study conducted across 5 healthcare facilities in Johannesburg by Abrams et al (2017), the median (95% CI) time to initiation from PCR test was 36 (31, 40) days and from receipt of results was 23 (18, 27) days [185].

2.4.2. Summary

Excluding the study which looked at ART initiation in only the first week after diagnosis, and the study based on interviews with caregivers (for which the reported coverage was considerably lower), the estimates of linkage to ART services were generally much higher in urban areas than in rural. Among the studies based at urban healthcare facilities in urban areas, estimates ranged between 61% and 100%, while at the only one based in a rural area, the estimate was slightly lower at 56%. The number of infants included was small for many studies, which would impact the reliability of their estimates.

Only three studies reported the median time to initiation of ART, with results varying from between 1 and 3 months, and was shorter in the more recent studies, which may reflect better understanding of and adherence to guidelines. No data on time to initiation were available from the studies which followed infants diagnosed at birth; initiation may be expected to occur more quickly in these settings, with fewer opportunities for loss-to-follow-up.

2.5. Viral suppression

2.5.1. Results

The number of studies identified for the review of viral suppression among infants initiating ART is shown in Figure 2.3. Of 781 papers identified following a search of PubMed, 672 were excluded following review of their title and abstract, with the most common reasons being not reporting virological outcomes (n=283) and not including infants (n=174). Of the remaining 109 papers, which were fully reviewed, 106 were excluded; 47 were excluded for not assessing virological outcomes, 44 for not disaggregating results by age, 9 because they covered the time period prior to the introduction of immediate ART for all infants in South Africa, and 6 for not including infants. In addition to the 3 remaining papers, 2 studies were identified from a review of other papers and 1 from a conference database, giving a total of 6 studies included in the review.

The characteristics of the 6 studies included are shown in Table 2.9. All but one of the studies used data recorded in patients notes, with the remaining study based on data from NHLS. Half

were based at a single healthcare facility and half across a group of healthcare facilities. Five studies were based exclusively in urban areas, with one using data from both urban and rural healthcare facilities.

Figure 2.3 - Flowchart of studies identified for the review of viral suppression (topic 3)
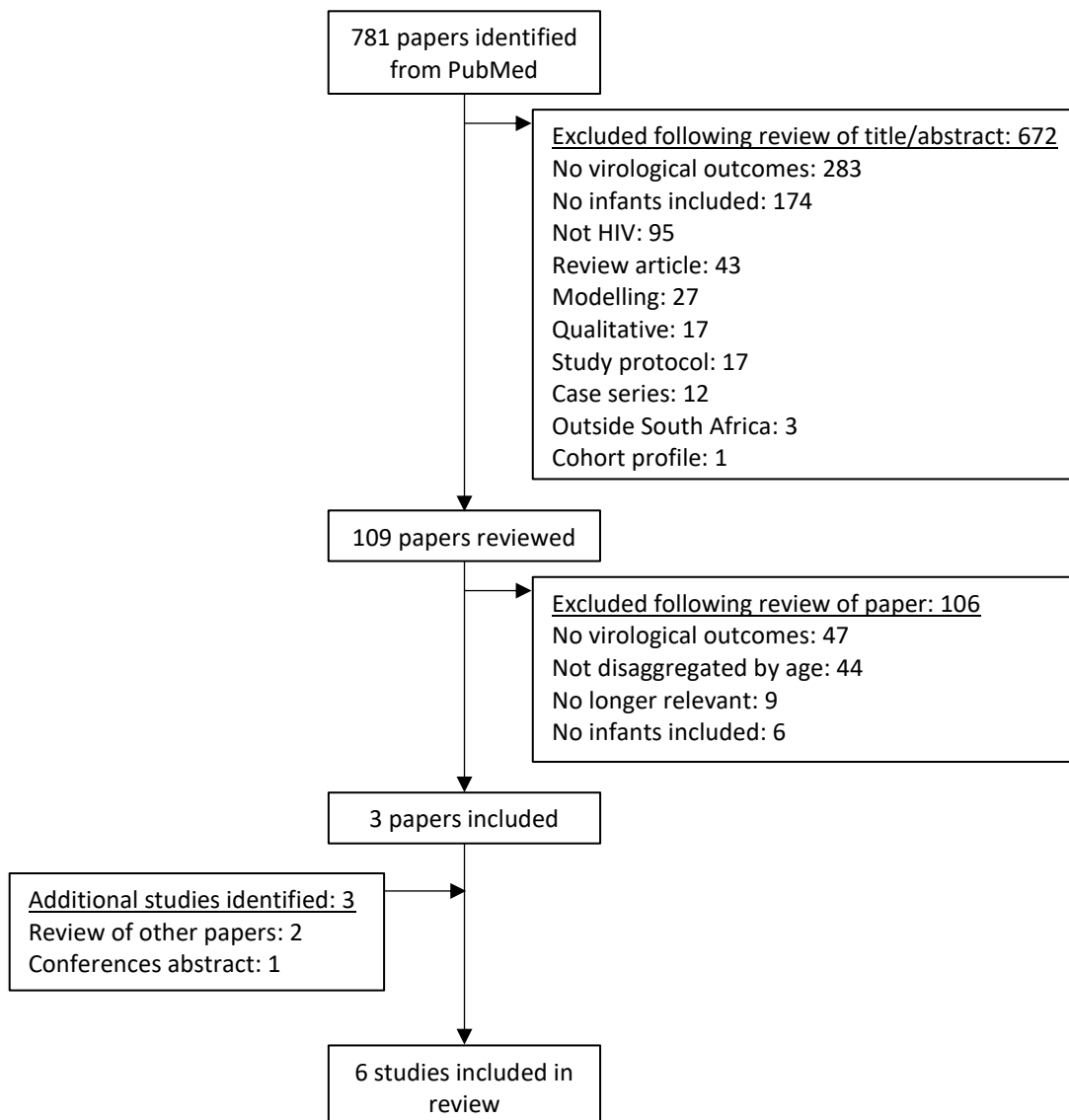
Table 2.9 - Characteristics of the 6 studies included in the review of viral suppression (topic 3)

	Studies reporting the proportion of diagnosed infants on ART achieving viral suppression
Total number of studies	6
Primary source of data	
Patient records	5
NHLS	1
Geographical coverage	
Single healthcare facility	3
Group of healthcare facilities	3
Setting	
Urban	5
Mixture of urban and rural	1

The results from the 6 studies are summarised in Table 2.10. All 6 studies reported the proportion of infants achieving viral suppression, however two studies did not report the threshold used, and among others various definitions were used (<50, <400, ≤400, <1,000 copies/mL). Viral suppression was assessed at multiple time points ranging from 3 to 24 months after ART initiation, with one study instead reporting time to viral suppression. Some studies reported the proportion of infants who had achieved viral suppression by a given time point after initiation (that is, the cumulative incidence of the proportion who had ever had suppressed viral load by that time, regardless of what subsequently happened), and some reported the proportion suppressed at a given time point (that is, the proportion suppressed at their closest measurement to a given time).

Five studies used data recorded in patient records [165, 177, 184, 187, 188]. Three studies, by Technau et al (2018), Teasdale et al (2018) and Dunning et al (2017), reported the proportion of infants suppressed by a given time point after ART initiation [165, 184, 187]. Technau et al (2018) included 72 infants diagnosed and initiated on ART at a single hospital in Johannesburg, of whom 56 were still in care one year later, by which time 40 (71%) had had a documented viral load ≤400 copies/mL [184]. Teasdale et al (2018) looked at viral suppression among 188 infants initiating ART at 5 healthcare facilities across the Eastern Cape, with the outcomes reported using two thresholds (<50 copies/mL and <1,000 copies/mL) and at three time points (6, 12 and 24 months after initiation of ART). The cumulative incidence of viral suppression increased with time from initiation, and was (by definition) higher using the higher threshold. By 24 months on ART, 61% and 76% of infants were suppressed <50 copies/mL and <1,000 copies/mL respectively [187]. The study by Dunning et al (2017) included 22 infants initiating ART at a single hospital in Cape Town, 18 had a viral load measurement available by 6 months on ART, of whom 28% were suppressed (with the threshold used not reported) [165].

Table 2.10 - Estimates of the proportion of infants initiating ART who achieved viral suppression

Study (First author, publication year)	Time period covered	Setting	Data source	Study design	Number of infants initiating ART	Definition (cps/mL)	Time since initiation	Number still in follow-up	Number and proportion suppressed, of those with a measurement
Patient records									
Technau, 2018 [184]	Sep 2013-Jun 2016	Single hospital, Johannesburg	Patient records	All infants diagnosed and initiated on ART at hospital	72	≤400	By 6m	Not reported	32/62 (52%)
							By 12m	56	40/56 (71%)
Teasdale, 2018 [187]	Apr 2013-Dec 2015	5 healthcare facilities, Eastern Cape	Patient records	All infants initiating ART <1 year at healthcare facilities included. Assessed cumulative incidence of viral suppression.	118	<50	By 6m	Not reported	n=77, 20% (95% CI 13, 28)
							By 12m	Not reported	n=88, 47% (95% CI 37, 56)
							By 24m	Not reported	n=42, 61% (95% CI 49, 71)
						<1,000	By 6m	Not reported	n=77, 42% (95% CI 33, 51)
							By 12m	Not reported	n=88, 68% (95% CI 58, 76)
							By 24m	Not reported	n=42, 76% (95% CI 66, 84)
Dunning, 2017 [165]	Jul 2013-Aug 2015	Single hospital, Cape Town	Patient records	All infants diagnosed and initiated on ART at hospital	22	Not reported	By 3m	18	3/18 (17%)
							By 6m	18	5/18 (28%)
Shiau, 2017 [188]	Jan- Dec 2011	3 hospitals and 2 clinics, Johannesburg	Patient records	All infants initiating ART aged <6 months	104	<50	At 6m	104	18/58 (31%)
Lilian, 2013 [177]	Aug 2008-Jul 2010	Single hospital, Johannesburg	Patient records	All infants born and diagnosed at hospital included. Follow up to December 2010.	23	<400	N/A	N/A	14/23 (54%), by median 18.4w (8.0, 64.3) after initiation
NHLS									
Moyo, 2018 [189]	Apr 2015-May 2016	Tshwane (urban) and uMkhanyakude (rural) districts	NHLS	Positive PCR tests in infants linked to viral load tests using automated patient linking algorithm	Not reported	Not reported	At 6m	88	39/75 (52%)

The fourth study based on patient records, by Shiau et al (2017), looked cross-sectionally at the proportion suppressed at 6 months after ART initiation (rather than the cumulative incidence) [188]. Among 104 infants who had initiated treatment at healthcare facilities across Johannesburg, reported viral suppression at 6 months was low at 31%, although only about half of infants in care had a measurement available. The fifth study, by Lilian et al (2013), was conducted at a single hospital in Johannesburg [177]. Among 23 infants initiating treatment, 14 (54%) achieved a viral load <400 copies/mL recorded, at a median 18.4 weeks after initiation.

The final study, by Moyo et al (2018), used data from the NHLS from across two sub-districts (one in an urban area and one in a rural area) [189]. Of 88 infants retained in care at 6 months of ART, 75 had an available viral load measurement, of whom 52% were suppressed.

2.5.2. Summary

Few studies looked at viral suppression, and all had relatively small sample sizes. As may be expected, there was high variation in the estimates of suppression according to both the definition used and the time point at which it was assessed. Only two studies looked at viral suppression cross-sectionally, rather than at the proportion ever suppressed by a certain time after ART initiation. The interpretation of the estimates was limited given that the majority of studies only looked at viral suppression within the first 6 months or 1 year on treatment. Among these studies, the proportion suppressed ranged from 17% to 71%. One study looked at suppression to 2 years after ART initiation, with 76% of infants ever achieving a viral load <1,000 copies/mL. In many studies, many infants who had initiated treatment had dropped out by the time point of interest or did not have an available measurement, further complicating interpretation.

2.6. Summary

There was high variation in estimates of retention across the cascade. Investigation of trends over time was made difficult by the use of different data sources and methods, although in individual studies that covered long time periods, retention across all stages generally improved.

The majority of studies were conducted in urban areas, with estimates from studies in rural areas or made at a national or provincial level, often lower. Healthcare facilities in urban areas may have better resources and fewer staff shortages, which would be expected to lead to higher estimates across the cascade stages. In addition, patients in rural areas may have to travel longer distances for care.

Many studies were conducted at a single healthcare facility, which may limit the generalisability of their findings. In addition, many of the studies were published by the same few healthcare facilities, which further limits the generalisability of the results described here to other settings.

Facilities with staff motivated to publish research may be more familiar with guidelines and recommendations, and thus would be expected to see higher than average retention across the cascade. There are several other limitations to the use of data from a single healthcare facility. Firstly, the inability to follow infants who move and receive care elsewhere would lead to the underestimation of each stage. Secondly, using data from only a single facility meant that many studies had relatively small sample sizes, especially for the ART and viral suppression stages, affecting the reliability of the corresponding estimates.

The high variation in estimates may partly by explained by the use of different methods and sources of data, each of which had different strengths and limitations. Using data from the NHLS enabled estimation of the cascade across larger regions. However, these studies required the use of data linkage to assess how many infants met each stage; the availability and quality of the data recorded and its impact on the reliability of linkage is unclear. In addition, for the testing stage, there was the additional challenge of having to deduplicate repeat tests on the same infant, which may have contributed to some estimates being greater than 100%. Many studies used data recorded in patient notes, which may be more reliable, although details on how data were recorded was rarely given. Estimates from studies that relied on other sources of data, such as interviewing caregivers or infants' road-to-health booklets, were often lower.

Only one other study has assessed the complete cascade among infants in South Africa [177]; among 838 infants exposed to HIV and born at a single hospital in Johannesburg between August 2008 and July 2010, 82% received a PCR test at 6 weeks of age, and of those diagnosed with HIV, 61% started ART, of whom 54% achieved viral suppression in a median 18 weeks. The use of data from the same population of infants across all stages improves the internal validity of results, making inference about likely movement of a single patient across the cascade easier. In addition, many of the studies which could only assess a single stage of the cascade relied on disparate data sources for estimating the numerator and denominator, in particular those looking at testing coverage, which may further affect their validity.

Chapter 3. Sources of data

3.1. Introduction

In this chapter, I introduce the Africa Health Research Institute (AHRI), and describe the three main sources of AHRI data which I will use to estimate the cascade of care throughout my thesis. I subsequently give a brief overview of several other sources of data which are also used.

3.2. AHRI

AHRI is a Wellcome Trust and Howard Hughes Medical Institute funded research centre based across two sites in the province of KwaZulu-Natal, South Africa. The institute was formed following the merger of two existing research centres, the Africa Centre for Health and Population Studies and the KwaZulu-Natal Research Institute for TB-HIV (K-RITH) in 2016. Population research is carried out at the Somkhele site, which lies within the Hlabisa health sub-district (shown in Figure 3.1).

Figure 3.1 - The Hlabisa health sub-district
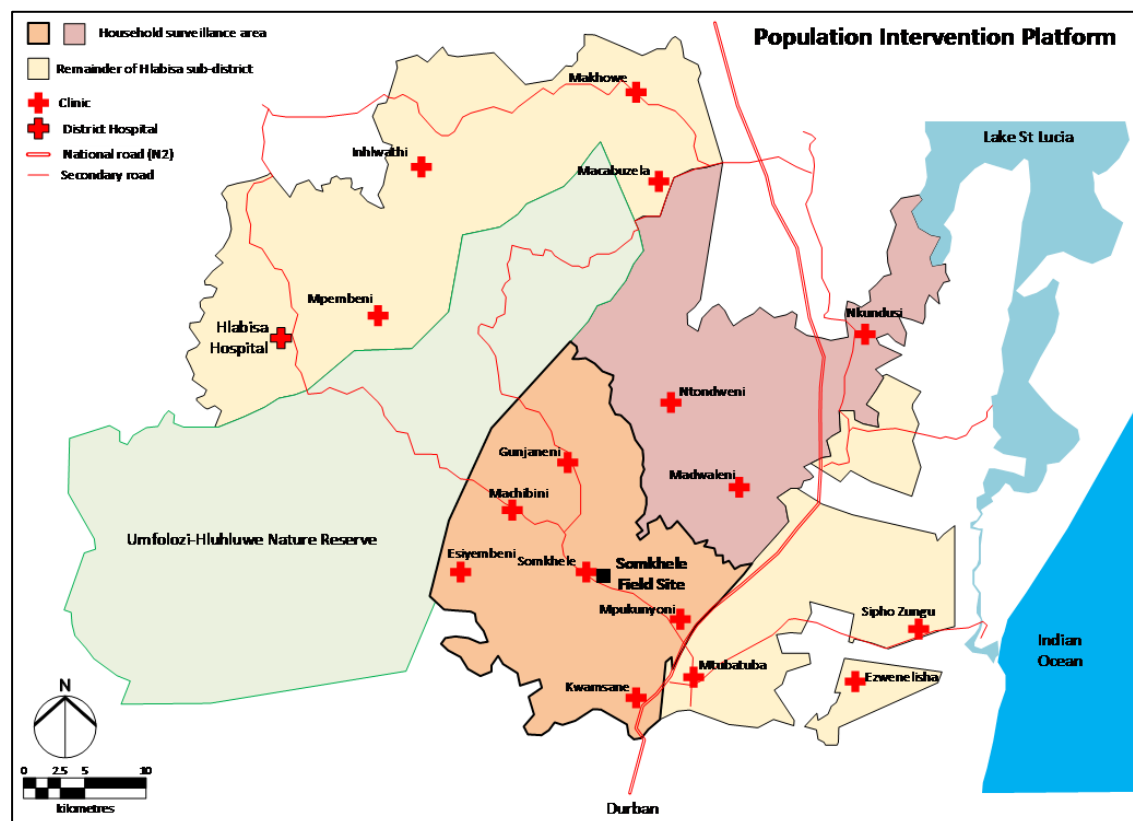
The region in orange represents the original AHRI demographic surveillance area, the region in pink the extension to the demographic surveillance area made at the beginning of 2017 (from which ACDIS data is not included in this thesis), and the region in yellow the remainder of the Hlabisa health sub-district.

3.3. AHRI data sources

In this thesis, three main sources of data from AHRI were used: demographic surveillance data (ACDIS, the Africa Centre Demographic Information System), an ART clinical database (ACCDB, the Africa Centre Clinical Database), and PCR test data from NHLS (National Health Laboratory Service). ACCDB and ACDIS are already linked by AHRI, but linkage between ACCDB and NHLS, ACDIS and NHLS, and deduplication of repeat PCR tests in the NHLS data will be conducted in Chapter 4 (Figure 3.2). Linkage was conducted from NHLS to both ACDIS and ACCDB, rather than for example simply linking to ACDIS and using the existing linkage between ACDIS and ACCDB to infer links between NHLS and ACCDB, so that any errors in the existing linkage were not compounded. Inconsistent links between the two methods were identified and resolved manually. Each of the datasets is described below.

Figure 3.2 - Links between ACDIS, ACCDB and NHLS PCR datasets
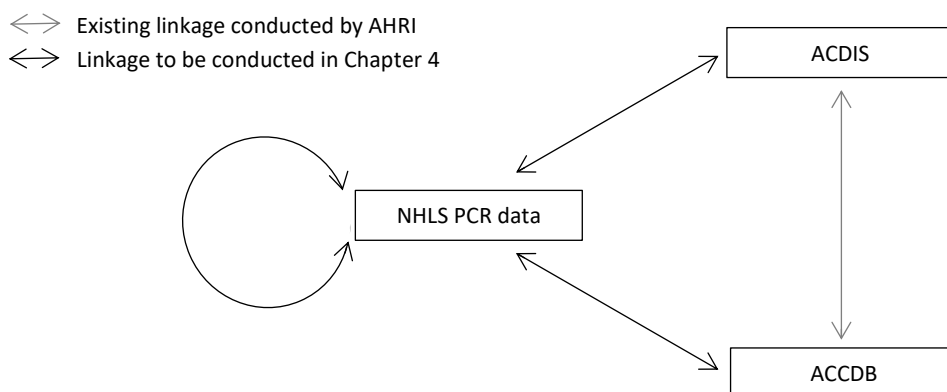
3.3.1. Africa Centre Demographic Information System (ACDIS)

The Population Intervention Platform (PIP) is a household demographic survey, which has been running in part of the Hlabisa health sub-district since 2000. Data are collected on all residents within the surveillance area by trained fieldworkers who visit each household annually, and include household membership, migration, pregnancies, deaths, socioeconomic data and childhood vaccination history (Table 3.1). The proportion of individuals who participate is over 99% [133]. Until the end of 2016 the surveillance area was that shown in orange in Figure 3.1, and covered a population of 85,000 people [133]. The surveillance area was extended at the beginning of 2017 to include the pink region, and now covers 845km^2 with a population of approximately 150,000 residents, however, the time period covered means that only individuals and data from the original surveillance area are included in this thesis. Data from the PIP demographic survey form ACDIS.

Since 2003, all residents and a sample of non-residents (migrants who return periodically to the household) aged ≥15 years have been offered a dried blood spot HIV test as part of the survey,

though results are not returned to individuals. Annual participation levels in the HIV surveillance component of the survey ranged from 26% to 46% between 2003 and 2012, although it was estimated that 77% of individuals had participated at least once within 9 years of becoming eligible [130].

The version of ACDIS used in this thesis included data to the end of 2016.

Table 3.1 - Description of the data available within ACDIS and their use within this thesis

Subject	Description of data collected	Use in this thesis
Individuals	Includes date of birth, sex, patients, household membership	
Socioeconomic survey	Includes data on education, employment, household infrastructure	Maternal socioeconomic factors explored as potential predictors of infant PCR testing in Chapter 6
Pregnancy	Delivery dates and outcome of all pregnancies	Identification of infants born to women with HIV
Migration	Describes all migration events within and in to/out of the surveillance area	Ascertain follow-up status of infants across cascade
HIV surveillance	All individuals ≥15 years offered a dried blood spot HIV test	Identification of women with HIV
Childhood vaccination history	Collected for children <5 years	Comparison of engagement in child health services with PCR testing coverage in Chapter 7
Hospitalisations	Data on all hospitalisation at Hlabisa hospital from the Hospital Information System (HIS)	Maternal hospitalisations explored as a potential predictor of infant PCR testing in Chapter 6
Deaths	Date, location and causes	Ascertain follow-up status of infants across cascade

3.3.2. Africa Centre Clinical Database (ACCDB)

Until the end of 2012, clinical data on individuals living with HIV were collected by AHRI employees stationed in each of the 17 healthcare facilities in the sub-district, and were stored in a database called ARTemis. ARTemis included HIV-related data from the patient's first visit at a clinic, including data prior to ART initiation.

ARTemis funding, through the President's Emergency Plan for AIDS Relief (PEPFAR), ended in 2012, so since 2013, AHRI has obtained ART data from TIER.net. TIER.net is the national ART surveillance system and contains individual-level data on individuals on ART at each clinic. A summary of the data available within ACCDB and their use within this thesis is shown in Table 3.2. At each clinic visit, the ART regimen prescribed, the length of the prescription given and results from recent laboratory tests (including viral load and CD4 measurements) are manually entered, based on the information recorded in each patient's paper notes. Importantly, only data after ART initiation are captured, and no data on ART prescribed for PMTCT are recorded.

Within the sub-district, TIER.net is used as an offline electronic register, meaning that the data held in the system in each of the clinics are not accessible from the other clinics. However, the system was designed as part of a three-tier approach to patient monitoring, with the first tier consisting of paper registers only and the third tier of a networked online electronic medical record system. The initial choice of system used within each facility is dependent on resources available, with the long-term aim that all facilities eventually transition to the fully networked system, with integration with laboratory systems and modules for collection of TB-related data and HIV-testing planned [190].

Table 3.2 - Description of the data available within ACCDB and their use within this thesis

Subject	Description of data collected	Use in this thesis
Individuals	Sex, date of birth, date of death	
ART data	Start and stop dates, regimens	Ascertain follow-up status, description of initial regimens, frequency of treatment interruptions in Chapter 8
Laboratory data	Includes CD4 and viral load measurements	Achievement of viral suppression, immunological status at ART initiation in Chapter 8

AHRI has a memorandum of agreement with the Department of Health to access TIER.net data from each of the healthcare facilities within the Hlabisa health sub-district. The memorandum covers linkage of individuals within TIER.net to other sources of data, with a waiver of the requirement for individual consent. Data from each of the clinics are pooled to form ACCDB. Given the offline nature of the current system, when data from the clinics are pooled, data linkage must be conducted to identify individuals who have transferred between clinics and thus have their ART history split over multiple (unlinked) records. Linkage is conducted by AHRI, and is based on variables including first name, surname, date of birth and South African national ID.

As only TIER.net data from clinics within the sub-district are available, a complete ART history is not available for individuals who initiated treatment elsewhere and subsequently transferred to a clinic in the sub-district. When clinics first enter data on an individual into TIER.net, they indicate whether the individual initiated ART at that clinic or was already on treatment, enabling the identification of these individuals within ACCDB.

The TIER.net system was only implemented in the sub-district in 2013, and so data prior to this were retrospectively entered by clinics. However, notes belonging to patients who had died or had been lost-to-follow-up before 2013 may not have been retained by the clinic, and thus may not have been back-captured into the system. In order to fill in this missing data, data (after ART initiation) on individuals in ARTemis meeting these criteria and not found in TIER.net were separately added into ACCDB.

Record linkage has also been used by AHRI to link individuals in ACDIS and ACCDB; South African national identifiers were used in conjunction with patient names, cell phone numbers, date of birth, sex, nearest clinic and location to link the individuals, using a deterministic and probabilistic algorithm.

The version of ACCDB used for analysis in this thesis contains follow-up until July 2017.

3.3.3. National Health Laboratory Service (NHLS) PCR data

All laboratory tests carried out in public facilities in South Africa (approximately 85% of all tests, with the rest conducted through private clinics or non-governmental organisations [191]) are processed through the NHLS. A form is completed (by hand in the case of facilities in the sub-district), with a unique barcode linking the specimen to this form, and both are sent to a central laboratory for processing. The results are returned to the clinic electronically.

All NHLS data are stored within a central data warehouse, and HIV DNA PCR test data for the 17 facilities within the Hlabisa health sub-district, as well as another clinic which lies just outside, were extracted by NHLS and sent to AHRI. Results were available for all tests from June 2010 to July 2017 for most clinics but, because of an extraction error, data from one clinic (which accounts for approximately 15% of all PCR tests done within the sub-district) were only available between January 2011 and March 2016. Variables available within this dataset include first name, surname, date of birth, sex, the clinic at which the test was carried out, and the individual's clinic ID. The standardisation of data items, for example the formatting of dates and coding of sex, is done by the NHLS data warehouse.

In the Western Cape, individuals are assigned a unique patient identifier, which allows record linkage between various health-related data systems including the NHLS database [192], however, no comparable identifier exists in KwaZulu-Natal or the other South African provinces.

Table 3.3 - Description of the data available within NHLS and their use within this thesis

Subject	Description of data collected	Use in this thesis
Individuals	Identifying variables including first name, surname, date of birth, sex, clinic ID, South African national ID	
PCR test data	Clinic at which conducted, date of test, test result	Frequency, timing and results of PCR testing in Chapters 5, 6, 7

3.4. Other sources of data

Throughout this thesis, I also use several other sources of data. A more detailed description of each of these is given in the relevant chapters, but an overview is shown in Table 3.4.

Table 3.4 - Summary of other data sources used throughout thesis

Data source	Description	Use in this thesis
MONARCH	Cluster randomised trial aiming to improve PMTCT processes within the sub-district	Estimation of PCR testing coverage in Chapter 5
SSA	Annual reports detailing the number of birth registrations, by both year of birth and district	Estimation of PCR testing coverage in Chapter 5
ANCHSS	Nationally representative survey of antenatal HIV seroprevalence	Estimation of PCR testing coverage in Chapter 5
DHIS	Monthly aggregated returns on activity collected on paper registers at each healthcare facility	Estimation of PCR testing coverage in Chapter 5

ANCHSS: National Antenatal Sentinel HIV & Syphilis Survey Report; DHIS: District Health Information System; MONARCH: Management and Optimization of Nutrition, Antenatal, Reproductive, Child Health and HIV Care; SSA: Statistics South Africa

Chapter 4. Data linkage

4.1. Introduction

In this chapter, individual-level PCR test data from NHLS, ART data from ACCDB, and data from ACDIS, the population demographic surveillance dataset, were linked to enable the construction of a cascade of care for HIV-exposed infants and infants with HIV in future chapters.

Given that it was expected that some infants would receive more than once PCR test, I firstly linked repeat tests conducted on the same infant within the NHLS database, in order to identify the number of infants with multiple tests and to create a dataset with one record per infant. These unique infants were subsequently linked to ACCDB, allowing the estimation of the uptake of ART in Chapter 8, and to ACDIS, for the estimation of the proportion of infants who received a PCR test, in Chapter 5. None of the datasets shared a unique patient identifier that would have allowed for simple matching of records on the same individual, and I therefore developed a deterministic and probabilistic linkage algorithm.

4.2. Objectives

The objectives of this chapter are to:

- Link PCR tests conducted on the same infant within the NHLS database
- Link infants with a PCR test to those reported to be on ART in the AHRI ART clinical database (ACCDB)
- Link infants with a PCR test to those in the population demographic surveillance database (ACDIS)

4.3. Methods

4.3.1. NHLS data cleaning

In total, 19,884 PCR test records conducted between June 2010 and July 2017 were extracted from the NHLS database by NHLS staff and sent to AHRI. In order to prepare the data for linkage, I then cleaned the data. During this process, a number of problems with the data were identified, which are described below.

It is known that if parents have not chosen a name for their baby by the time of the infant's PCR test then the mother's name is often recorded on the request form instead. For some records this was clearly indicated, for example the name was reported as "Baby of [mother's name]", although it is likely that for others it was not, meaning there was no way of identifying those with the incorrect name recorded. This may make the linkage of repeat tests on the same infant

difficult. This was accounted for in the development of the linkage algorithm, as described below.

First name and surname were sometimes reversed. In cases where the probability of what was recorded in the first name field being a first name as opposed to a surname was less than 10% (based on an existing reference dataset of first name and surnames derived from the AHRI population database, ACDIS), and equivalently the probability that what was reported in the surname field was unlikely to actually be a surname was less than 10%, the two names were swapped.

Sex was missing for approximately 25% (5,156/19,884) of records. It was decided to impute sex in situations where the probability of each name being male, or female (again, as calculated from the AHRI reference dataset), was >80%. Given the known problem with recording mother's name as the first name, there was concern that unknown sex may incorrectly be imputed as female for some records. To investigate this, the sex distribution within the dataset was summarised after imputation, and found to be 53.3% female and 46.7% male. Although this differed from national estimates of the sex ratio at birth from Statistics South Africa (49.5% female and 50.5% male) [193], there was little difference from the ratio for those with sex completed in the original dataset (52.3% female and 47.7% male) and so was considered acceptable. Sex was imputed for 45% (2,332/5,156) of records where it was missing.

The reported clinic ID often changed between tests for records which otherwise appeared to refer to the same individual, with the recorded ID often partially based on the date of the test. This was accounted for in the linkage algorithm.

The NHLS test request form asks for South African national ID but this is often not recorded, as it is not a required field. This is particularly true for PCR test data, as infants are not assigned an ID number until after the civil registration of their birth has taken place. National ID was therefore missing for more than 99% of records, and so could not be used as a linking variable.

4.3.2. Linkage algorithm

The data linkage process consisted of the following three stages:

- Stage 1: Linkage was conducted within the NHLS dataset to identify repeat tests on the same individual. Estimation of the number of infants tested and of the number of infants with multiple tests was important to enable the accurate assessment of clinician adherence to infant testing guidelines and of the proportion of infants who received a positive test result. In addition, the creation of a dataset containing one record per infant simplified the matching process for the subsequent stages.

The newly identified unique individuals from the NHLS PCR dataset were then linked to the following datasets:

- Stage 2: Link infants in the NHLS PCR dataset to ACCDB, the AHRI HIV clinical database
- Stage 3: Link infants in the NHLS PCR dataset to ACDIS, the AHRI population database

4.3.2.1. Linkage algorithm for stage 1

There are two main types of data linkage method, deterministic and probabilistic. Deterministic linkage identifies and matches records with exact agreement on a defined set of variables, and is the preferred choice when data are of high quality or share a unique identifying variable (for example, national ID) on which to match [194]. In probabilistic linkage, all records are pairwise compared and a similarity score is assigned to each pair according to the level of agreement on each of the matching variables, with the decision about whether or not to consider each pair a true match made based on this score. Probabilistic methods work substantially better than deterministic methods, even in datasets where there are low rates of data entry error [195]. Data linkage algorithms typically consist of several successive steps, and in this case it was decided to initially identify exact matches using a deterministic step (in order to reduce the number of pairwise comparisons subsequently required), before using a main probabilistic matching step.

Step 1 - Deterministic linkage

Identify records matching exactly on the following six key linkage variables:

- first name
- surname
- date of birth
- sex
- the name of the clinic at which the PCR test took place
- the patient's clinic ID

Step 2 - Probabilistic linkage

Following the deterministic step, additional links were identified using probabilistic linkage based on the same six linkage variables.

One technique often used in data linkage, to reduce the number of pairwise comparisons made, is blocking, whereby the dataset is partitioned into mutually exclusive and exhaustive groups based on the value of a blocking variable [196]. Pairwise comparisons are then only made between records within the same group. For example, if year of birth was used as a blocking variable, only records identifying individuals born in the same year would be compared. This

means links would never be made between records with different values of the blocking variable. Although multiple iterations of probabilistic linkage steps could be made with a different choice of blocking variable each time to avoid missed links due to data errors, the variables chosen should still be well and reliably completed. The choice should also take into account the number of possible values of the blocking variable, as blocking is most valuable (in terms of the reduction in the number of pairwise comparisons to be made) when the variable used partitions the dataset into many different groups. For example, sex is rarely used as a blocking variable. Given the characteristics of the linkage variables available within the NHLS datasets and the data issues with each described above, blocking was not felt to be appropriate here.

Multiple methods for calculating the similarity score required during probabilistic linkage exist. Rule-based methods (for example, +1 to the score if the fields match exactly, +0.5 if the fields differ by only one character, +0 otherwise) are the easiest to implement, but values are usually chosen arbitrarily and can lack statistical justification [197]. Instead, a score based on the probability of agreement was used here. This method was first proposed by Newcombe and Kennedy in 1962 [198], but was later formalised by Fellegi and Sunter [199]. Here, an agreement weight and a disagreement weight for each matching variable is calculated as $log_2 \frac{m}{u}$ and $log_2 \frac{1-m}{1-u}$ respectively, where m is the probability that two records on the same individual match on that variable and u is the probability that two records on different individuals match on the variable. The m-probability encompasses the possibility of data entry error or of true values changing over time (for example in this setting, clinic, as infants may receive tests at more than one clinic), and the u-probability represents the likelihood of chance agreement. If the two records match on a given field the agreement weight (which is always positive) contributes to the overall similarity score, otherwise the disagreement weight (which is always negative) does. If one or both of the values for a field are missing or unknown, a weight of zero contributes to the score. The final similarity score is calculated as the sum of the field weights across all linkage variables, with higher scores signifying records that are more similar.

To estimate the u-probabilities, a dataset of unique infants was required [197]. To do this, all tests taken on the same day as the recorded date of birth were selected (n=2,246), making the assumption that no child received more than one test on the same day. Those taken on the date of birth were chosen, rather than on any other particular day, as this gave the highest number of tests from which to estimate the probabilities. From this so called "birth cohort", the probability of chance agreement for each of the fields was estimated, with the final values shown in Table 4.1. The u-probabilities of chance agreement on first name, surname, clinic and clinic ID were estimated at 0.0033, 0.0120, 0.1339 and 0.01 respectively. The u-probabilities of the day, month and year of birth and of sex were calculated as 1 divided by the number of

Table 4.1 - Estimated m- and u- probabilities and the corresponding agreement and disagreement field weights for each variable

		m probability - probability that identifier agrees in a true matching pair		u probability - probability that identifier agrees in a non-matching pair		$w_a =$ agreement field weight = $\log_2 \frac{m}{u}$	$w_d =$ disagreement field weight = $\log_2 \frac{1-m}{1-u}$
		Notes	m	Notes	u		
First name*		Mother's name may be used for early tests	0.8	Estimated using "birth cohort"	0.0033	7.921	-2.317
Surname*			0.9	Estimated using "birth cohort"	0.0120	6.229	-3.305
Date of birth	Day		0.95	1/30 days in a month	0.0333	4.847	-4.274
	Month		0.95	1/12 months in a year	0.0833	3.512	-4.196
	Year		0.95	1/8 years of study	0.125	2.926	-4.129
Sex			0.95	1/2	0.5	0.926	-3.322
Clinic		Can move clinics	0.7	Estimated using "birth cohort"	0.1339	2.386	-1.530
Clinic ID		Often unknown, appears to change with calendar year for some clinics	0.4	Estimated using "birth cohort"	0.01	5.322	-0.722
						Maximum score = 34.069	Minimum score = -23.795

Note: A field weight of 0 was used if one or both of the values was missing or unknown
*In order to account for partial agreement, the field weights for first name and surname were adjusted using the scaled Levenshtein edit distance, see Table 4.2.

possible values, and were 0.0333 (=1/30), 0.0833 (=1/12), 0.125 (=1/8) and 0.5 (=1/2) respectively.

As no equivalent dataset of known matches was available to estimate the m-probabilities, the final values were chosen from a review of published literature [194, 200, 201], accounting for the known problems with the variables in the NHLS dataset described above. The values chosen are shown in Table 4.1. The m-probability for surname was 0.9, to reflect the potential for spelling mistakes or transcription errors to be made. First name had a lower m-probability (0.8); given the known issue with recording of mother's name, it was possible that an early test conducted on an infant would be requested using their mother's name but subsequent tests would be requested under their own name. Sex and the variables for the date of birth were given relatively high m-probabilities (each 0.95) to reflect the fact that these variables do not change over time and that data entry errors would be unlikely. Although a mistake in the recording of clinic would be very unlikely, infants may transfer between clinics and so this variable was given an m-probability of 0.7. Because of concern over the reliability of clinic ID, its m-probability was low at 0.4.

The corresponding agreement and disagreement weights for each field, and their relationship to the m- and u- probabilities, can also be seen in in Table 4.1. For example, for sex, each individual would be expected to match with approximately half of all other records on this field but data entry errors are unlikely. The chosen m- and u- probabilities give the desired result that sex contributes a relatively low agreement weight (0.926) to the similarity score if pairs match, but the penalisation is relatively high for records which do not match (-3.322).

The Fellegi-Sunter method only allows a binary classification of two values into either 'agree' or 'disagree', but it may be desirable to account for partial agreement in the case of some variables. This is particularly true for first name and surname as, for example, "Lizzie" and "Lizzy" are more likely to refer to the same individual than "Lizzie" and "Alice", and the weights used in each case should reflect this. There are several possible methods for making this adjustment. For instance, a cut-off could be used to assign the agreement field weight to records that are a relatively close match as well as those matching exactly; however, this method has the disadvantage that the same weight is used for both exact and partial matches. Alternatively, multiple weights could be used and assigned, depending on the extent of agreement between records [202]. This was the method chosen here for first name and surname.

Here, the extent of agreement was determined based on the value of an approximate string comparator algorithm, the Levenshtein edit distance [203]. The Levenshtein edit distance between two strings is defined as the number of single-character edits (additions, deletions and substitutions) required to turn one string into another, and can be scaled to a value between 0

and 1 by dividing by the length of the longer string, an upper bound on the distance [204]. Initially a linear mapping from (0, 1) to (disagreement weight, agreement weight) was considered [203], however following review of the corresponding scores this was seen to result in a significant penalisation to the field weight of pairs where names were judged to be relatively similar. Instead, an alternative mapping which better represented the relationship between increasingly different pairs of names and the change in Levenshtein distance was chosen empirically, as shown in Figure 4.1. Example names giving a range of different scaled Levenshtein distance and their corresponding field weights under the two methods are shown in Table 4.2. For instance, under the linear mapping, the names "SIPHETHELO" and "SPETHELO" would receive a score a whole point less than two records with identical first names, while under the alternative mapping the difference was only 0.17. Similarly, two names judged as very unlikely to refer to the same individual (such as "MATHULI" and "SIMHLELELE") were given a weight far from the disagreement weight under the linear mapping (a difference of 2.05 in the example given), and this was corrected with the alternative mapping (to a difference of 0.07).

Figure 4.1 - Mapping of scaled Levenshtein distance to field weight for first name
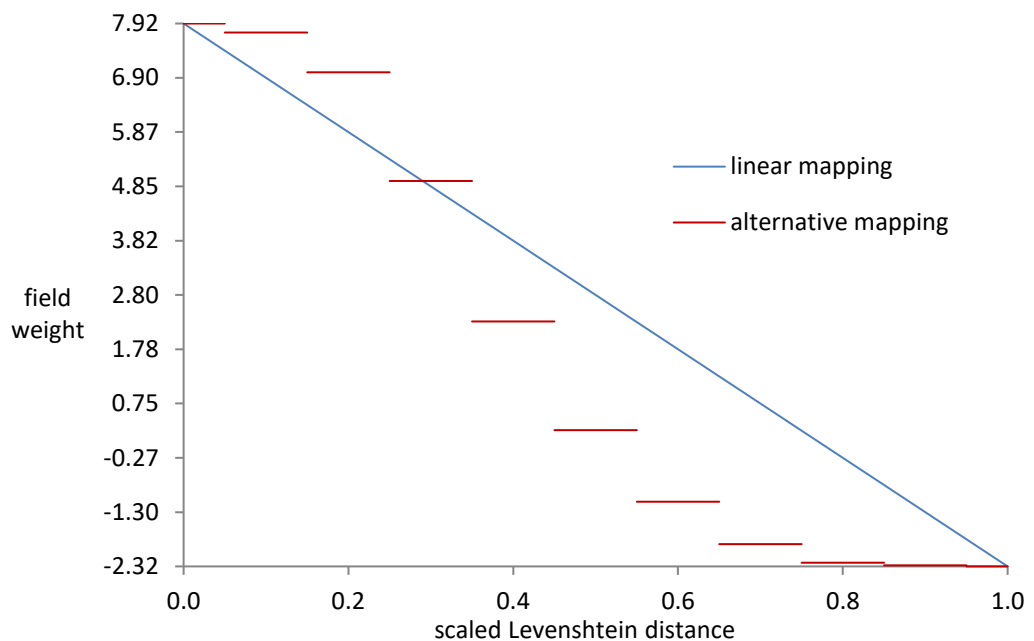
Table 4.2 - Adjusted field weights for first name, using both a linear mapping and an alternative non-linear mapping.

Scaled Levenshtein distance	Field weight, by mapping method		Example names	
	Linear	Alternative		
0.00 - 0.05	7.92	7.92	AYANDA	AYANDA
0.05 - 0.15	6.90	7.75	SIPHETHELO	SPHETHELO
0.15 - 0.25	5.87	7.00	LUVOLWETHU	VOLWETHU
0.25 - 0.35	4.85	4.95	SIPHINDILE	SPHILILE
0.35 - 0.45	3.83	2.30	NONDALO	NONKUTHALO
0.45 - 0.55	2.80	0.25	ABONGWE	ANOTHIWE
0.55 - 0.65	1.78	-1.10	NELIZWE	SAMUKELIWE
0.65 - 0.75	0.75	-1.90	SUMSILE	MASWAMAHLE
0.75 - 0.85	-0.27	-2.25	MATHULI	SIMHLELELE
0.85 - 0.95	-1.29	-2.30	KHANYISANI	MAWETHU
0.95 - 1.00	-2.32	-2.32	LIGUGU	SNOTHILE

Partial agreement on date of birth was accounted for by considering the day, month and year of birth as separate fields.

An example of the calculation of the final similarity score for two records is shown in Table 4.3. This pair of records match exactly on surname, month and year of birth, sex and clinic name, and therefore the agreement weight contributes to the final score for these fields. The two first names have a scaled Levenshtein distance of 0.17, and a weight of 7.00 is therefore used. The records have different days of birth recorded and thus here the disagreement weight is used. Finally, as the clinic ID was missing for record 1, a zero weight contributes for this field. The final similarity score is therefore 18.705.

Table 4.3 - An example calculation of the similarity score for two NHLS PCR records

		Agreement weight	Disagreement weight	Record 1	Record 2	Agree? (scaled Levenshtein distance)	Weight
First name		7.921	-2.317	Anhele	Anele	N (0.17)	7.000
Surname		6.229	-3.305	Dlamini	Dlamini	Y (0.00)	6.229
Date of birth	Day	4.847	-4.274	23	29	N	-4.274
	Month	3.512	-4.196	December	December	Y	3.512
	Year	2.926	-4.129	2013	2013	Y	2.926
Sex		0.926	-3.322	F	F	Y	0.926
Clinic		2.386	-1.530	KwaMsane	KwaMsane	Y	2.386
Clinic ID		5.322	-0.722	-	562/13	-	0

Total weight = 18.705

Once the similarity score had been calculated for all pairs, each was then classified as either a 'potential match' (defined as any pair with a score above a certain threshold) or 'not a match'. Potential links were manually reviewed to decide whether or not they should be considered a 'true match'. A histogram showing the distribution of the similarity scores is shown in Figure 4.2. The threshold for manual review was chosen as 15, both based on this distribution and by visual inspection of the data, and the choice was a trade-off between limiting the number of pairs to be manually reviewed to a feasible number and capturing as many true matches as possible.

Figure 4.2 - Histogram of similarity scores
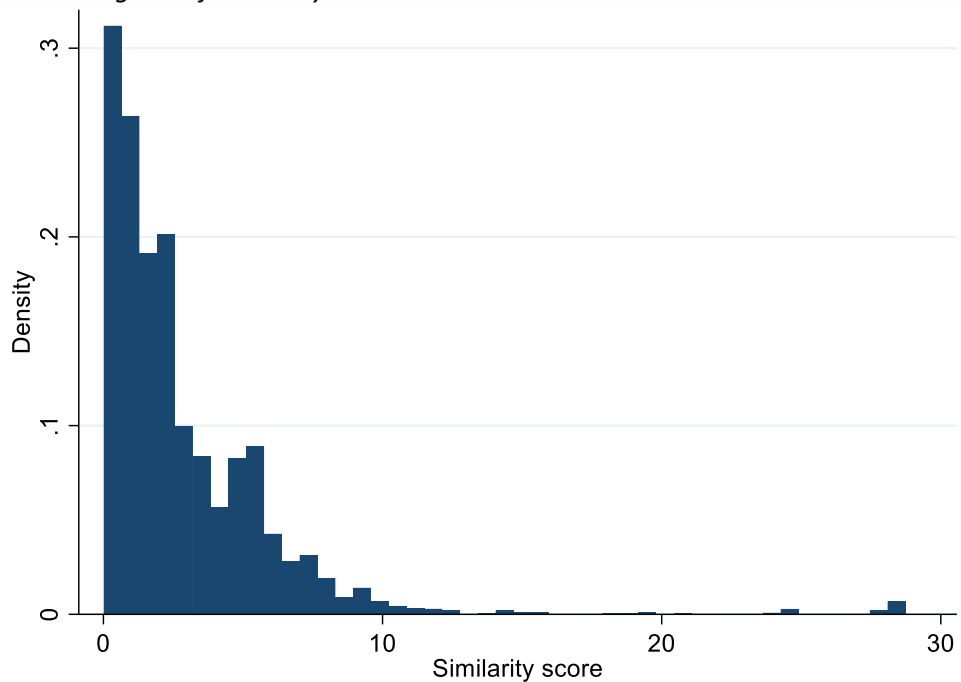
For clarity, only similarity scores above 0 are shown in the histogram.

As individuals may receive more than two PCR tests, it was expected that any given record may be a 'potential match' with more than one other record. Potential links were therefore grouped into 'networks', with each network reviewed together. This is demonstrated in Figure 4.3 where, for example, records A, B, C, and G have all been grouped into a single network based on high similarity scores (indicated by a 1 in the table) between A and B, A and G, B and G, and C and G. In this example, record C has been included in the network because of its similarity score with record G, despite the fact it was not deemed 'similar' to either record A or B. Following review, not all records within the network have to be linked together and considered to belong to the same individual. For example, records A, B and G could be linked, with record C considered as a separate individual.

Figure 4.3 - An example of records grouped into networks

	A	B	C	D	E	F	G	H
A		1	0	0	0	0	1	0
B			0	0	0	0	1	0
C				0	0	0	1	0
D					0	0	0	0
E						1	0	0
F							0	1
G								0
H								
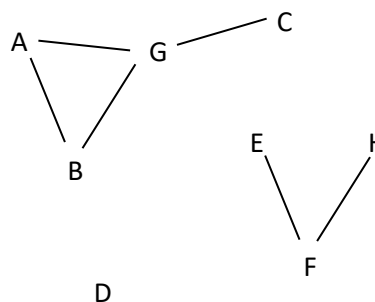
Note: 1 represents a similarity score greater than the threshold (i.e. a potential match), and 0 represents a similarity score less than the threshold.

Along with the values of the six linkage variables and their corresponding field weights, additional data fields were taken into account during the manual review process. This included the date of each PCR test (as an individual would be unlikely to be given two PCR tests close together at the same clinic), the result of each PCR test (because it would be unlikely for an infant to receive a negative test followed by a positive one), and how common both the first name and surname are, based on AHRI's existing name reference database as previously mentioned (as chance agreement is less likely with uncommon names).

Once all records to be linked had been identified, any discrepant values had to be reconciled in order to create a single set of identifying variables for each individual that could be used in the subsequent matching stages. The following rules were used sequentially to select the final value for each field.

- First name:
  1. Ignore names reported for tests in the first month of life
  2. Select the most common name within the network
  3. Select the most common name in the whole population
  4. Select a value at random
- Surname:
  1. Select the most common name within the network
  2. Select the most common name in the whole population
  3. Select a value at random

- Date of birth:
    1. Ignore dates after any test date
    2. Select the most common date within the network
    3. Choose date that results in tests timed in accordance with the South African national PCR testing guidelines (which recommended tests at 6 weeks of age prior to April 2015, and at birth and 10 weeks thereafter)
    4. Select a value at random
- Sex:
    1. Select the most common sex within the network
    2. Select the most likely sex for their first name
    3. Select a value at random
- Clinic:
    1. Select the clinic at which the most recent test was taken

As ART treatment data were not available from the one clinic outside of the sub-district from which PCR test data were included, those individuals who were only tested at this clinic were dropped from the analysis at this stage, making the assumption that they were born outside of the sub-district. In addition, those born before June 2010 (for whom early tests may not be captured) or after December 2016 (to allow 6 months for the receipt of all recommended tests) were excluded.

4.3.2.2. Linkage algorithm for stages 2 and 3

In stages 2 and 3, each unique infant identified in stage 1 was compared to each infant in ACCDB and ACDIS respectively. As both ACDIS and ACCDB contain data on individuals of all ages, only those born since June 2010 were selected for consideration in order to reduce the number of comparisons required, despite the risk of infants truly born in the time period of interest but with their date of birth incorrectly recorded being missed. The algorithms used for these stages were broadly similar to that used for stage 1, apart from the following adaptions.

As clinic ID was not available in either ACDIS or ACCDB, this was not used as a matching variable in stage 2 or 3. Furthermore, clinic name was also not available in ACDIS and so was not used in stage 3. The same agreement and disagreement weights were used for the remaining variables. Given the smaller number of variables available here for matching, the deterministic step was removed so that all linked records were manually reviewed.

Only one-to-one matches were expected here, that is, at most one infant from the PCR dataset should link to each infant in ACCDB or ACDIS, and vice versa. Potential matches were therefore again reviewed in networks, in order to ensure that multiple matches were not made.

Several additional variables were used for inspection during the manual review process including date of ART start (stage 2), mother's and father's name, and the last date at which the mother was known to be HIV negative (stage 3).

4.3.3. Assessing the quality of the data linkage

Validation of the results of record linkage usually involves the calculation of measures such as sensitivity, specificity and the error rate [205, 206]. However, these require a 'gold-standard' linked dataset, which was not available here.

To assess the quality of the linkage in stage 1, the proportion of infants in the matched dataset who had the same date of birth, sex and surname was calculated. This was compared to the same proportion in the "birth cohort", as a reference for agreement due to coincidence or multiple births. Date of birth, sex and surname were chosen as they were considered the most reliably completed variables. No alternative method or source of data was available to assess the linkage from stages 2 or 3.

4.4. Results

The results from stage 1 of the data linkage are shown in Figure 4.4. There were 19,884 PCR test records extracted from the NHLS database. In the deterministic step, 71 pairs were matched exactly on the 6 linkage variables, and after linking these, 19,813 potentially unique individuals remained. In the probabilistic step, each of these was pairwise compared to each of the others, resulting in approximately 200 million comparisons being made. 5,701 records were found to be a potential match to at least one other record (with a similarity score >15), and from these potential matches 2,658 networks were formed. 2,344 networks contained only 2 records, and, following manual review, 378 of these pairs of tests were judged to belong to 2 different infants and 1,966 were judged to belong to the same infant, resulting in a total of 2,722 unique infants. 251 of the networks contained 3 records, 55 contained 4 records and 8 contained 5 records, and from these 344, 100, and 20 unique infants were identified respectively. Including the 14,112 records not found to be a potential match to any others, a total of 17,298 unique infants with a PCR test were identified. After applying the inclusion criteria requiring infants to have received at least one test at a clinic inside the sub-district and to have been born between June 2010 and December 2016, 15,234 infants remained.

Of these 15,234 infants, 5.7% had the same date of birth, sex and surname as another infant, compared to 4.0% in the "birth cohort" of unique infants.

The results from stage 2 of the linkage are shown in Figure 4.5. Each of the 15,234 infants with a PCR test were compared to each of the 727 individuals born since June 2010 on ART in ACCDB, resulting in 11,075,118 comparisons being made. From these, 388 potential matches were

identified, which formed 345 networks. Following manual review, 283 true matches were identified.

The results of stages 3 are shown in Figure 4.6. There were 15,996 individuals in ACDIS born since June 2010 in ACDIS. Following 266,877,260 pairwise comparisons with the 15,234 infants with as PCR test, 2,809 potential matches were identified and manually reviewed. This resulted in 2,349 (15%) infants with a PCR test being linked to ACDIS.

Figure 4.4 - Results from data linkage stage 1
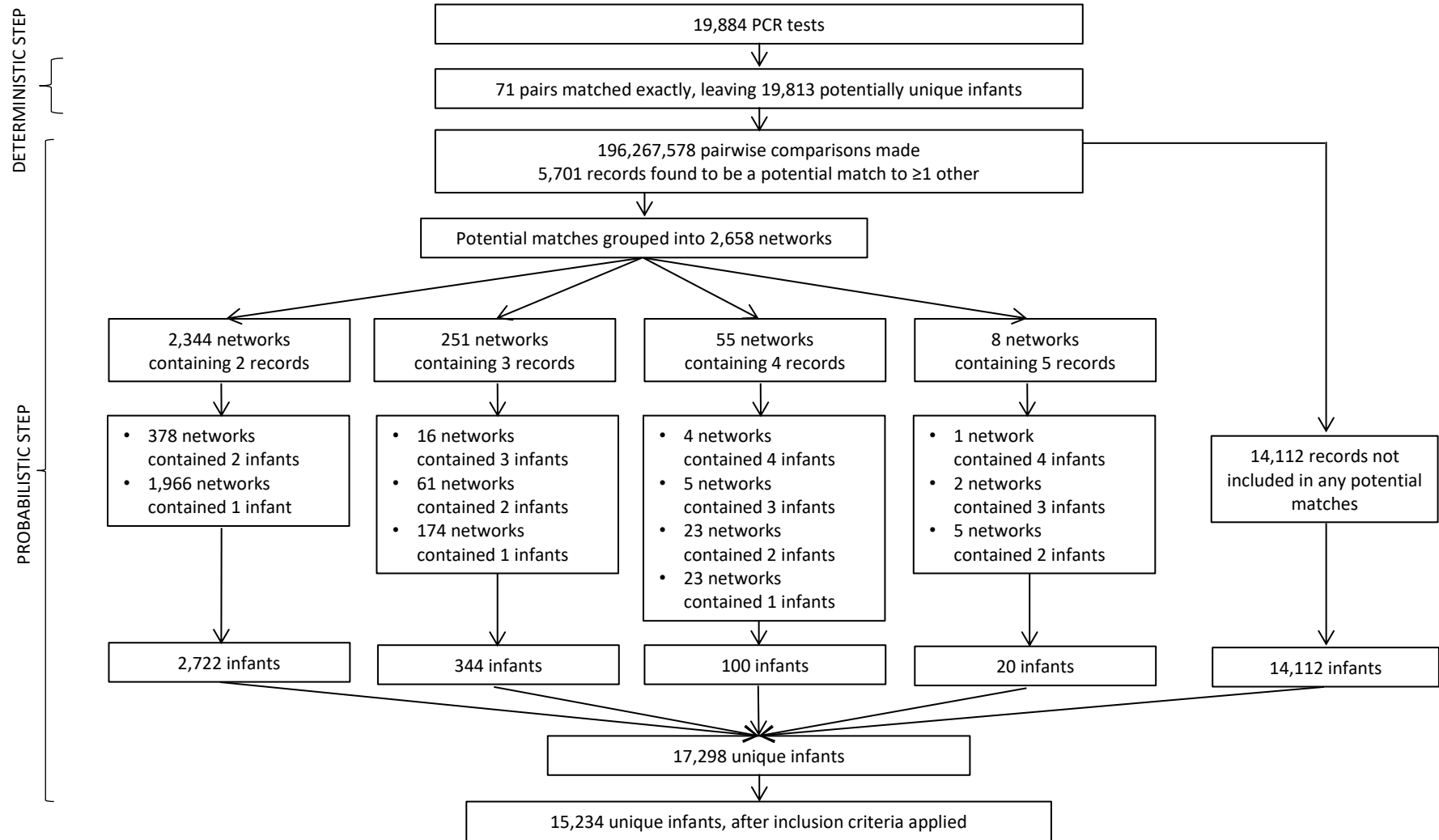
Figure 4.5 - Results from data linkage stage 2
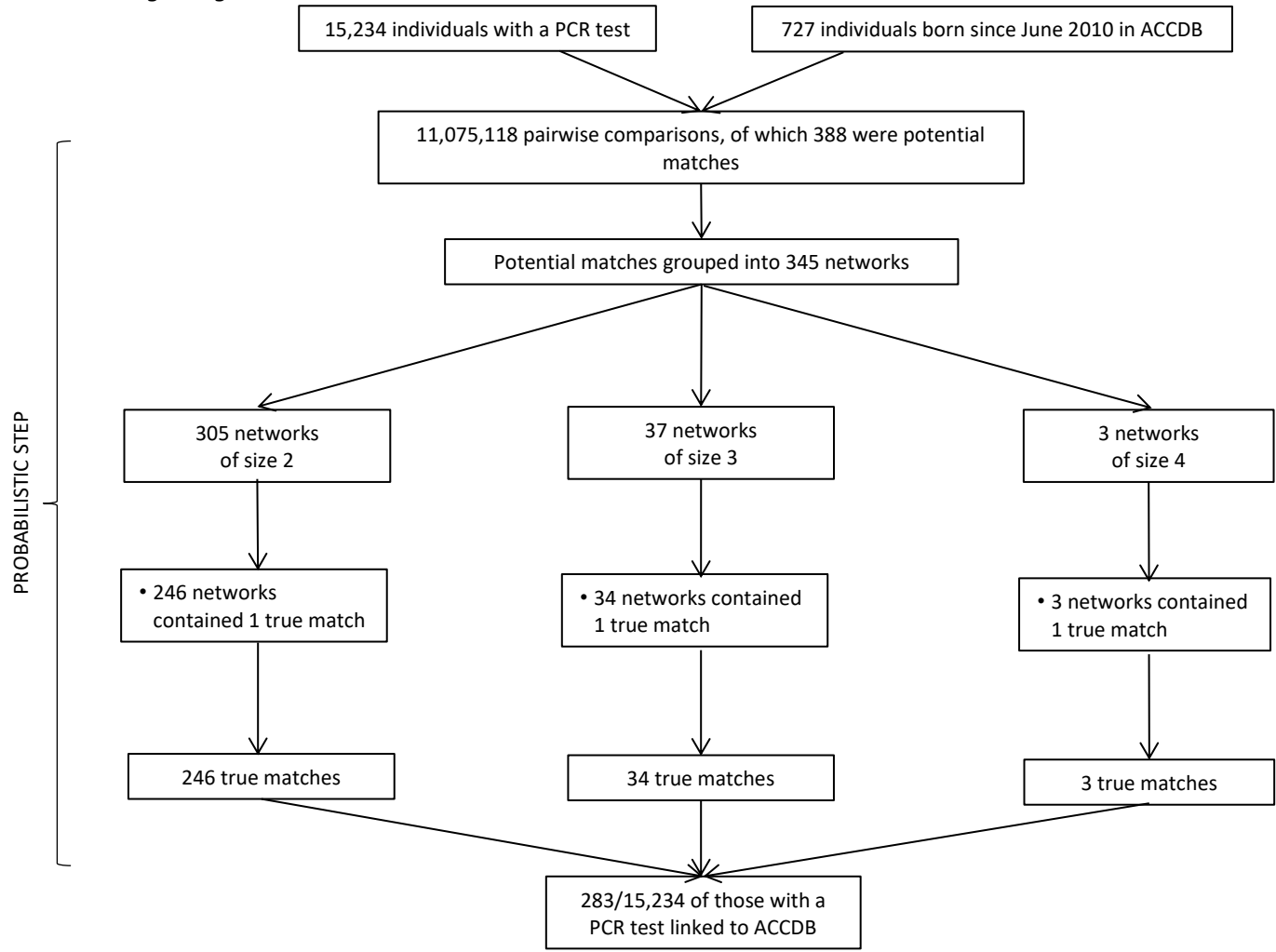
Figure 4.6 - Results from data linkage stage 3
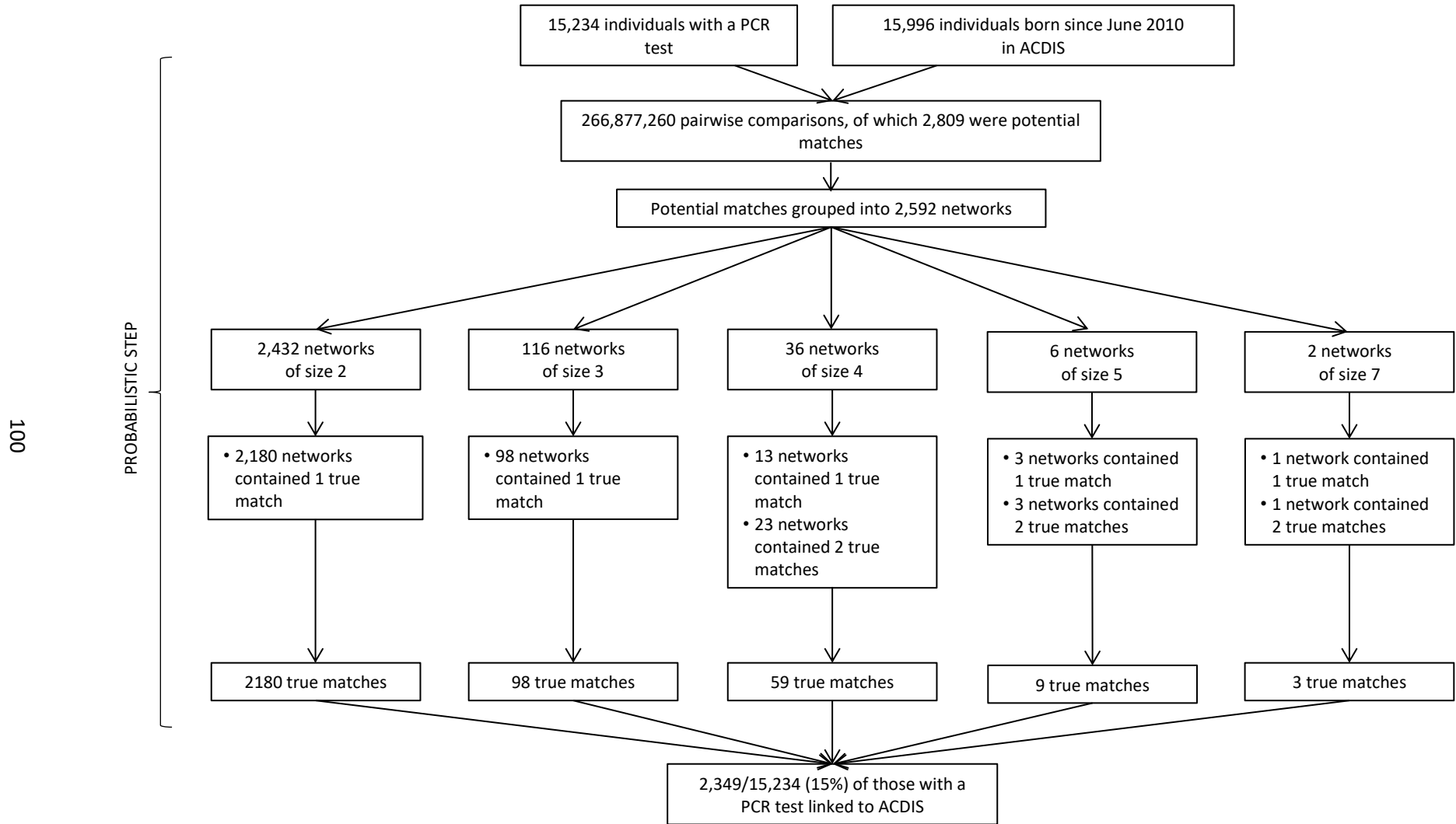
4.5. Discussion

Using a data linkage algorithm consisting of both deterministic and probabilistic steps, I was able to identify unique individuals with PCR tests recorded in the NHLS database, and subsequently link these individuals to both ART treatment and population demographic surveillance datasets. This enabled the construction of a cascade of care for HIV-exposed infants in subsequent chapters.

The results from the assessment of the quality of the stage 1 linkage suggested that the algorithm performed reasonably well, with only a slightly higher proportion of infants matching on key identifying variables than would be expected. Given the lack of unique identifiers available, as well as data issues with the reporting of names for newborn infants in particular, it is possible that some links may have been missed.

There are a number of additions that could have been made to the linkage algorithm, but these were not included. Firstly, the use of the Levenshtein edit distance to measure agreement between names would only capture some types of data entry error and misspellings. Soundex [207] is a phonetic coding system often used in record linkage to match names which sound similar (rather than just those which are spelt similarly), but was designed for names pronounced in English and its ability to match isiZulu names is unclear. Secondly, the use of value-specific rather than global u-probabilities could have improved the algorithm by accounting for the fact that, for example, some surnames are less common than others, and as such are less likely to match by chance. Thirdly, some matches may have been missed as a result of only considering individuals in ACDIS and ACCDB who were born since 2010 in stages 2 and 3. Finally, other more accurate methods could have been used to improve the values of the m-probabilities, such as Expectation-Maximisation [208], but were not used due to lack of time.

The algorithm developed for this analysis could also be applied to link the results of other laboratory tests stored in NHLS data, perhaps most usefully CD4 and viral load measurements, to the AHRI datasets. Although some of the data quality issues seen with the PCR data, for example the poor reporting of national ID and of first name, are likely to be less problematic with these data, the increased numbers of records and individuals would make this matching practically very computationally intensive as well as increasing the likelihood of overlap in the matching variables.

4.6. Key findings

The key findings from this chapter are:

- 19,884 PCR tests were extracted from NHLS database, from which 15,234 unique infants (who met the inclusion criteria) were identified. Of these, 283 infants were linked to ACCDB, and 2,349 were linked to ACDIS

- There were few linkage variables available for each step, for example South African national ID was poorly recorded in the NHLS dataset, and there were issues with the recording of the variables which were available, for example, the use of mother's name instead of the infant's name on the test request form

- There were few ways to validate or assess the performance of linkage algorithm, although a comparison of the proportion of infants in the deduplicated NHLS database with the same surname, date of birth and sex as another infant with the proportion that would be expected due to chance agreement, suggested some underlinkage

Chapter 5. Infant HIV PCR testing coverage

5.1. Introduction

In this chapter, I calculate and compare estimates of the coverage of PCR testing among HIV-exposed infants from four different methods, each based on several different sources of routinely collected data. I subsequently discuss the limitations to each of these methods, and the expected effect of these limitations on the resulting estimates.

5.2. Objectives

The objectives for this chapter are to:

- Explore the use of different methods in estimating overall HIV PCR testing coverage in HIV-exposed infants (that is, the proportion who ever had a PCR test) for the Hlabisa sub-district for June 2010 to December 2016, using a variety of surveillance and research data sources.
- Compare coverage by year of birth, and before and after the introduction of birth testing to the South African national guidelines in April 2015.

5.3. Methods

The following four methods were used to estimate HIV PCR testing coverage (Table 5.1):

- Method 1 (NHLS-AHRI surveillance): The numerator was the number of infants with a linked PCR test, among the denominator of infants born to women with HIV within the AHRI demographic surveillance area (identified through both the ACDIS and ACCDB datasets), following the linkage conducted in Chapter 4
- Method 2 (NHLS-SSA/ANCHSS): The numerator was the number of infants who received a PCR test according to the NHLS PCR data (following deduplication of repeat tests in Chapter 4), and the denominator was the number of HIV-exposed infants born which was calculated from estimates of the number of live births from Statistics South Africa (SSA) and of antenatal seroprevalence from the National Antenatal Sentinel HIV & Syphilis Survey Report (ANCHSS)
- Method 3 (NHLS-DHIS): The numerator was the number of infants who received a PCR test according to the NHLS PCR data (following deduplication of repeat tests in Chapter 4), and the denominator was the number of HIV-exposed infants which was estimated using District Health Information System (DHIS) data
- Method 4 (Road-to-health booklets): Using data routinely captured in the Road-to-health-booklets which were photographed as part of the MONARCH (Management and Optimization of Nutrition, Antenatal, Reproductive, Child Health and HIV Care) cluster-

randomised trial, with the numerator being the number of infants with a test recorded and the denominator the infants whose booklet was photographed.

Table 5.1 - Summary of the four methods used to estimate HIV PCR testing coverage

	Method 1: NHLS-AHRI surveillance	Method 2: NHLS-SSA/ANCHSS	Method 3: NHLS-DHIS	Method 4: Road-to-health booklets (MONARCH)
Numerator	Number of infants with a linked NHLS PCR test	Number of infants with a PCR test in the NHLS database	Number of infants with a PCR test in the NHLS database	Number of infants with a PCR test recorded in their Road-to-health booklet
Denominator	Infants born to women testing positive in serosurvey or initiating ART in ACCDB	Number of live births from SSA multiplied by antenatal HIV seroprevalence from ANCHSS	Aggregated returns of number of live births to women with HIV sent from clinics to DHIS	Infants born to women with HIV attending antenatal care at one of 7 demographic surveillance area clinics*
Linked numerator and denominator?	Yes	No	No	Yes
Geographical area covered	AHRI demographic surveillance area	Hlabisa sub-district	Hlabisa sub-district	AHRI demographic surveillance area
Time period covered (infant's date of birth)	June 2010 to December 2016	June 2010 to December 2016	April 2014 to December 2016	July 2015 to December 2016

*There are only 6 clinics within the AHRI demographic surveillance area, but a 7th which lies just outside in a more urban area was also included in the trial as patients often travel further to receive care here.

A comparison of the time periods covered by each of the methods is shown in Figure 5.1. NHLS PCR data were available for infants born between June 2010 and December 2016, so it was possible to estimate coverage for this entire time period using the first two methods, however DHIS data (used for method 3) were only available from April 2014 onwards. NHLS data from one clinic in the sub-district were only available for the period from January 2011 to March 2016, meaning only infants born between January 2011 and September 2015 would be expected to have data on all their recommended PCR tests. Data from the MONARCH trial (used for method 4) covered infants born between July 2015 to December 2016.

The proportion of infants who ever received a PCR test was estimated for the whole time period, by calendar year of the infant's birth, and separately before and after the introduction of birth PCR testing into the South African national guidelines (that is, for those born before 1st April 2015 vs. after 1st April 2015), with comparisons across time periods made using chi-squared tests.

A comparison of the geographical regions covered by the testing coverage estimates from each of the methods is shown in Figure 5.2. Using methods 2 and 3 is was possible to estimate testing coverage for the entire Hlabisa sub-district, but with methods 1 and 4 it was only possible to estimate coverage within the (smaller) AHRI demographic surveillance area.

Figure 5.1 - Comparison of the time periods covered by each of the four methods
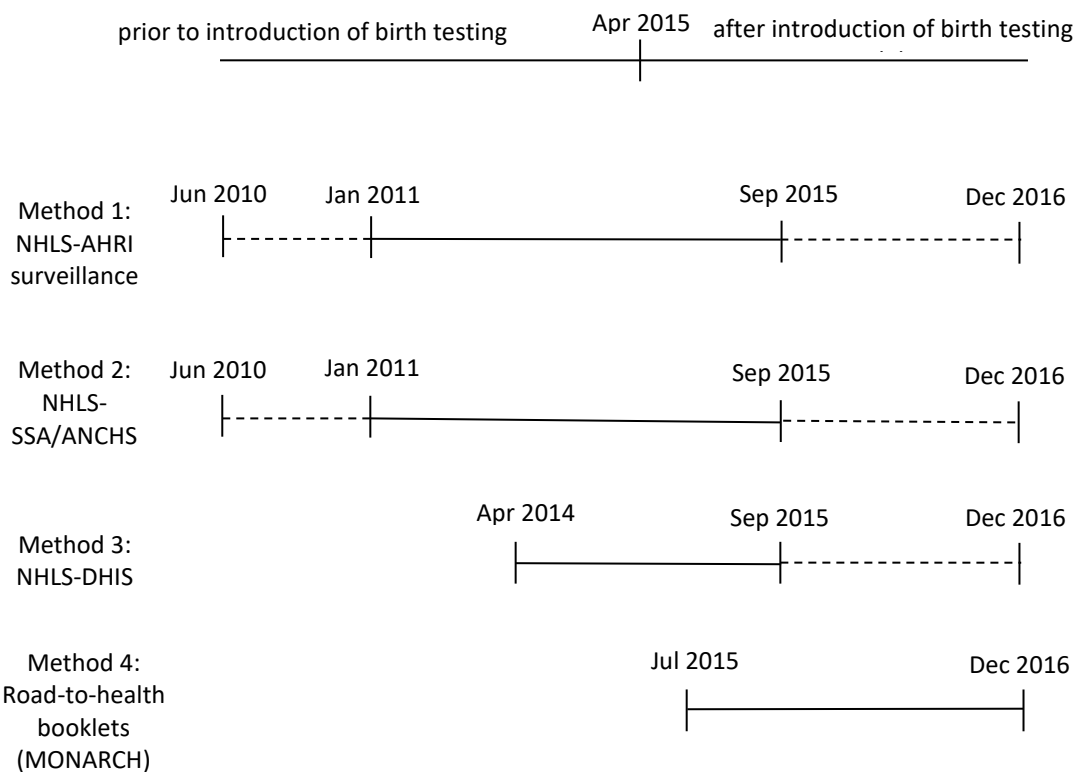
Dates presented refer to infant's date of birth. NHLS data (which is used in methods 1-3 is not available for infants born between June 2010 and January 2011, and September 2015 and December 2016 for clinic 6 (indicated by the dashed line).

Figure 5.2 - Comparison of the geographical area and size of population covered by each of the four methods
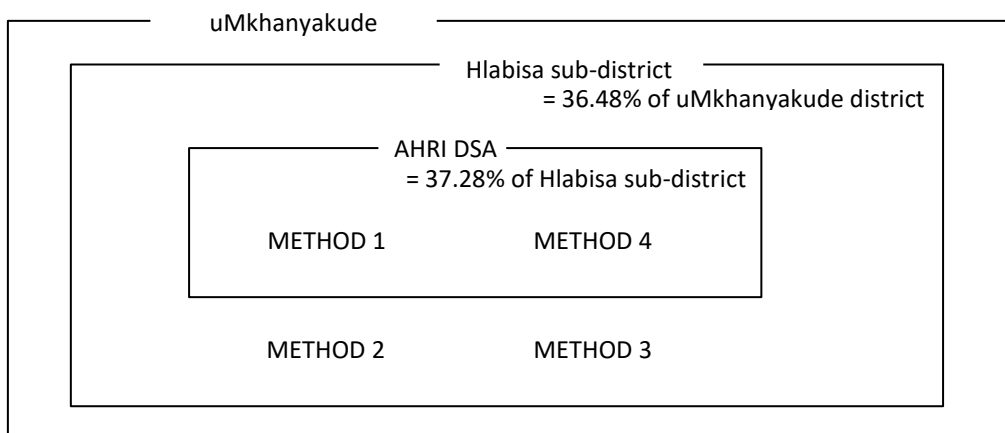
DSA: Demographic Surveillance Area. The population of the uMkhanyakude district is 625,000 [209], of the Hlabisa sub-district is 228,000 [132], and of the AHRI demographic surveillance area is 85,000 [133].

Using the first three methods, it was possible to estimate the proportion of exposed infants who were ever tested, however since the MONARCH trial only collected data on infants until they attended their 6 week post-natal visit, only coverage to this time point could be estimated for method 4. I therefore additionally calculated coverage to 7 weeks of age (as less than a quarter of infants in MONARCH had had their 6 week visit by 6 weeks of age, increasing to 90% by 7 weeks of age) using the first three methods, to enable comparison to method 4.

Method 1 - NHLS-AHRI surveillance

The first method used the results of data linkage between infants born to women with HIV within the demographic surveillance area and the NHLS PCR test data in Chapter 4 to estimate the proportion of infants tested.

Women living in the demographic surveillance area (who participate in ACDIS) can be identified as having HIV either because they tested positive in the annual serosurvey (with data available in ACDIS) or because they are recorded as being on ART in ACCDB (which has already been linked to ACDIS by AHRI). The date of HIV infection for each woman was estimated as either the date they first tested positive in the serosurvey, or the date of ART initiation, or the earlier of these two dates if they were identified using both methods. Infants born to these women are identified during annual ACDIS data collection, and those who were born between 1st June 2010 and 31st December 2016, were born after the assumed date of maternal HIV infection, and were known to be resident within the demographic surveillance area at birth (and so would be expected to receive their PCR tests at one of the clinics in the sub-district) were included in the analysis, giving the denominator of HIV-exposed infants.

Individuals in ACDIS were linked to NHLS PCR test data in Chapter 4 (stage 3) to estimate the numerator, so from this testing coverage could be estimated as the proportion of the identified HIV-exposed infants with a linked test.

Method 2 - NHLS-SSA/ANCHSS

For method 2 the testing coverage was estimated using the number of infants who ever received a PCR in the Hlabisa sub-district according to the NHLS PCR data, divided by an estimate of the number of HIV-exposed infants born within the sub-district. The number of exposed infants (the denominator) was calculated using estimates of the number of live births from SSA and of the antenatal HIV seroprevalence from ANCHSS.

SSA publishes annual reports detailing the number of live births, by both year of birth and district, with figures originating from the number of registered births [193, 210-215]. In South Africa it is a legal requirement that all births are registered, and this should be done within the first 30 days of life. Late registration of births outside of this timeframe is allowed but additional

supporting documentation, for example a certificate from the facility at which the child was born, is required. Reports containing the number of registered births for each year are published in the middle of the following year (e.g. the 2012 report, which is published in the summer of 2013, contains the number of live births registered for each year of birth, to the end of 2012). With each new report released, the number of births for each of the previous years increases, incorporating late registrations. Thus the published figures will underestimate the true number of infants born in a given year, particularly for more recent calendar years, where a greater number of births will not yet have been registered. To account for this, I calculated the average percentage increase with each year based on numbers reported in previous reports, up to 5 years after the initial reported estimate, to adjust the most recently reported estimate for likely future late registrations.

Estimates of the antenatal seroprevalence were taken from ANCHSS [216]. ANCHSS is a national sample, in which a sample of 1,595 public health facilities across South Africa (all of which have at least 20 first time antenatal clinic attendees per month and have the ability to do HIV testing) is taken. An HIV antibody test is conducted on all women aged 15-49 years attending antenatal care at each clinic over a given time period, following verbal informed consent. This test is conducted in addition to normal routine care and so all women, even those already known to have HIV, are included. From this, the seroprevalence of HIV can be estimated. Prior to 2015, only women attending for their first antenatal care visit at the sample clinics were included in the survey, but the design was changed from 2015 onwards allowing all women to be included, although each woman was never included more than once. Although the survey is repeated annually, only estimates to 2015 were available by the time of my analysis, and so the prevalence in 2016 was assumed to be the same as in 2015. Seroprevalence is reported at both a national and district level, with, for example in the 2015 survey, the estimate for the uMkhanyakude district, within which the Hlabisa sub-district lies, based on 361 women from 20 clinics (equating to an average of 18 women per clinic).

Women who acquired HIV during pregnancy but after they were tested for the survey would not be reflected in the ANCHSS estimates, but as South African guidelines recommend HIV testing of all women at delivery they should still be diagnosed before the birth of their infant, and their infant would therefore still be expected to receive a PCR test. To correct for this potential bias, I adjusted the ANCHSS estimates of seroprevalence to include the expected number of additional incident HIV infections that would have occured during pregnancy. According to a cohort of 25,608 women who were pregnant within the sub-district between 2010 and 2015, the rate of incident HIV during pregnancy is estimated to be 4.5 (95% CI 3.4, 5.8) per 100 person-years [148], corresponding to 0.0008654 (95% CI 0.0006538, 0.0011154) per person-week. In order to calculate the risk of HIV infection between sampling in ANCHSS and delivery, data from

the MONARCH trial were used to estimate the distribution of gestational age at antenatal care visits; the mean gestational age at first antenatal care visit (19.5 weeks) was used for 2010 to 2014, and the mean gestational age at one randomly selected visit per woman (37.5 weeks) was used for 2015.

To calculate the number of HIV-exposed infants born within uMkhanyakude each year, the estimate of the number of live births was multiplied by the adjusted antenatal seroprevalence. The population of the Hlabisa sub-district is estimated to be 228,000 [132], and of the uMkhanyakude district is estimated to be 625,000 [209]; these estimates were used to scale down the number of infants born within the whole uMkhanyakude district to the number born within the Hlabisa sub-district, making the assumption that the distribution of live births and of HIV prevalence was uniform across the district.

For comparison, testing coverage based on unadjusted estimates of the number of live births, unadjusted estimates of antenatal seroprevalence, and PCR test data prior to linkage of repeat tests (that is, assuming each test conducted was on a unique infant), was also calculated.

Method 3 - NHLS-DHIS

As with method 2, method 3 used the number of infants tested in the sub-district according to NHLS PCR data as the numerator, but instead used data from DHIS to estimate the number of HIV-exposed infants born for the denominator.

All public health facilities in South Africa send aggregated data on their activity, including the number of live births to women with HIV, on paper registers to DHIS. These data are then entered into an electronic database by DHIS, with monthly reports compiled at a clinic, district and national level. The denominator was estimated as the sum of the number of infants born in all the healthcare facilities in the sub-district, adjusted to account for infants not born in a healthcare facility, estimated using ACDIS to be 3% of all births. For comparison, the proportion of infants tested based on the unadjusted estimates of the number of live births, and PCR test data prior to linkage of repeat tests was calculated.

Method 4 - Road-to-health booklets (MONARCH trial)

Method 4 used data from Road-to-health booklets photographed in the MONARCH trial. MONARCH was a stepped wedge cluster randomised trial that captured data on all women attending antenatal care at one of the seven clinics in the AHRI demographic surveillance area between July 2015 and January 2017. The trial was designed to evaluate the impact of a continuous quality improvement intervention package on PMTCT processes. Continuous quality improvement is the "repetitive cycle of process and outcome measurement, design and implementation of interventions to improve the processes of care, and re-measurement to

determine the effect on quality of care" [217], and the MONARCH trial used tools such as process maps and plan-do-study-act cycles [218]. The trial had two primary outcomes: (i) the proportion of pregnant women with HIV with an antenatal viral load test, and (ii) the proportion of pregnant women initially presenting without HIV who received at least one repeat HIV test. Following a baseline data collection phase, the intervention was sequentially rolled out to clinics in 2 month phases (called steps). Results from the trial to date have shown a positive impact of the intervention on increasing the proportion of women with HIV with an antenatal viral load test, but not on repeat HIV screening among women initially testing negative [138].

Maternal case records, which are a handheld maternity medical record including information on HIV testing from which HIV status can be determined, are retained by clinics following delivery. As these records are considered routine Department of Health data, they could be photographed by MONARCH trial team staff without consent being required from women. If a woman who had been seen for antenatal care at a clinic within the demographic surveillance area delivered at a clinic outside of the region, then their case record would not be available. Women were also asked to participate in interviews about their antenatal care experience, and whether the Road-to-health booklets of their infants could be photographed, for which informed consent was required. This occurred at delivery, the 3-6 day postnatal visit, and the 6 week visit. The booklet is issued to all infants born in South Africa and includes records of HIV PCR testing, vaccinations and growth measurements throughout childhood [219]. Data from the photographs of the maternal case records and the Road-to-health booklet were entered into an electronic dataset by MONARCH trial staff.

Testing coverage was estimated as the proportion of infants whose Road-to-health booklet was photographed that had a PCR test recorded. In order to assess the completeness of recording of PCR tests within Road-to-health booklets, the number of PCR tests reported to be conducted at the seven clinics within the demographic surveillance area according to NHLS was compared to the number reported in the MONARCH data. Linkage between MONARCH and ACDIS datasets, which would have enabled direct comparison of the differences in NHLS and MONARCH PCR test data at an individual level, was planned by AHRI staff but had not been completed by the time of my analysis.

5.4. Results

Method 1 - NHLS-AHRI surveillance

A total of 2,254 HIV-exposed infants were identified in ACDIS, of whom 1,314 (58%) were born to a mother identified as having HIV through the serosurvey and linkage to ACCDB, 631 (28%) were born to a mother identified only in the serosurvey, and 309 (14%) were born to a mother

Table 5.2 - Method 1: Estimates of the proportion of HIV-exposed infants who received a PCR test based on linkage between NHLS and AHRI surveillance data, overall, by year of birth and by guideline time period

Guideline time period	Calendar year of birth	Number of HIV-exposed infants identified	Number with a linked PCR test	Number with a linked PCR test by 7 weeks of age	**Proportion of infants ever tested**	**Proportion of infants tested by 7 weeks of age**
	Source of data	AHRI surveillance	NHLS	NHLS		
	Method of calculation	A	B	C	**B/A**	**C/A**
Prior to introduction of birth testing	June - December 2010	192	73	37	**38%**	**19%**
	2011	386	183	107	**47%**	**28%**
	2012	342	164	93	**48%**	**27%**
	2013	349	167	94	**48%**	**27%**
	2014	437	187	126	**43%**	**29%**
	January - March 2015	121	48	39	**40%**	**32%**
	Total	1,827	822	496	**45%**	**27%**
After introduction of birth testing	April - December 2015	240	89	44	**37%**	**18%**
	2016	187	54	24	**29%**	**13%**
	Total	427	143	68	**33%**	**16%**
TOTAL		2,254	965	564	**43%**	**25%**

identified by linkage to ACCDB only. The proportion of infants with a linked PCR test is shown in Table 5.2. Overall, 965 (43%) infants ever received a PCR test, and this proportion varied by calendar year from 29% in 2016 to 48% in 2012 and 2013, with no discernible trend year on year but with a higher proportion ever tested among those born prior to the introduction of birth testing compared to after (45% vs. 33%, p<0.001). The proportion tested by 7 weeks of age ranged from 13% among those born in 2016 to 32% between January and March 2015.

There was no statistically significant difference in the proportion of infants ever tested according to the method of identification of HIV status of the mother; among infants born to mothers identified though both the serosurvey and ACCDB, the serosurvey only, and ACCDB only, the proportion of infants tested was 44%, 40% and 41% respectively (p=0.167).

Method 2 - NHLS-SSA/ANCHSS

SSA data on the number of live births in uMkhanyakude, by year of birth, and for each available published report between 2010 and 2016, are presented in Table 5.3. The numbers are also shown graphically in Figure 5.3. Reported numbers for each year of birth were available from the year itself onwards, though updated numbers for 2010 and 2011 were not published in the 2016 report. Initial estimates of the annual number of births increased steadily with each year from 10,592 in 2010 to 13,858 in 2016. The estimated number of births in 2012 dropped from 16,381 in the 2014 report to 14,706 in the 2015 report, with no explanation given. All other estimates increased year on year.

Table 5.3 - Method 2: Number of live births in uMkhanyakude, by year of birth and year of report

		Year of report						
		2010	2011	2012	2013	2014	2015	2016
	2010	10,592	12,673	13,201	13,650	13,911	14,013	
	2011		11,107	13,421	14,315	14,722	14,785	
Calendar year of birth	2012			11,722	15,744	16,381	14,706	18,088
	2013				13,231	16,398	16,689	17,474
	2014					13,353	15,774	16,738
	2015						12,992	16,340
	2016							13,858

Note: Year of report refers to the most recent year with data included in the report. For example, the 2012 report, which is published in the summer of 2013, contains estimates of the number of live births per year to the end of 2012

Figure 5.3 - Method 2: Number of live births in uMkhanyakude, by year of birth and year of report
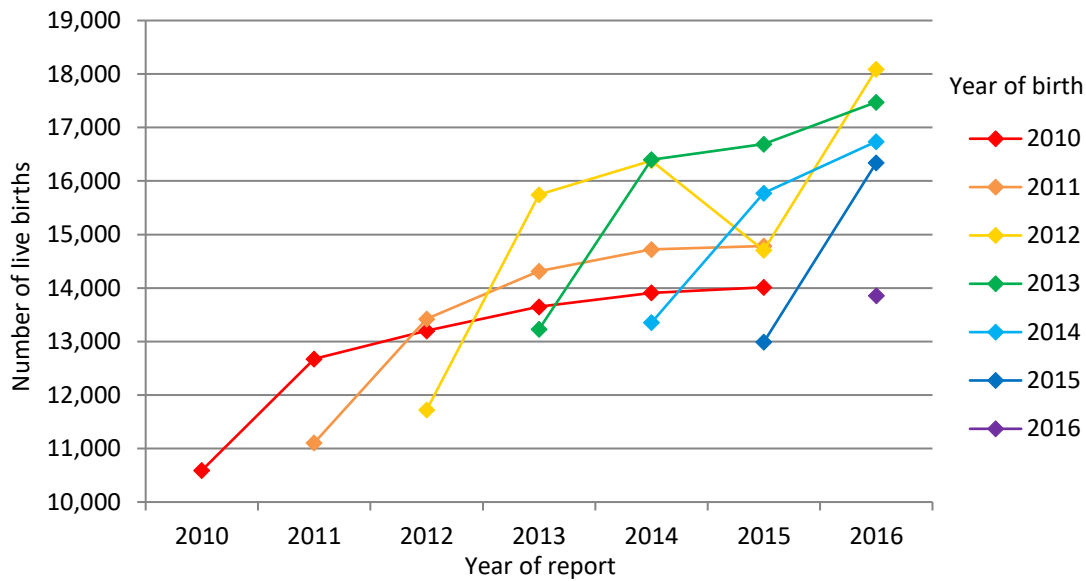
For each year of birth, the percentage increase in the estimate of the number of births with each new report (as compared to the initial estimate reported) is shown in Table 5.4 and Figure 5.4. For example, the initial estimate of the number of births in 2011 (as reported in the 2011 report) was 11,107, and by the 2014 report (3 years later) this had increased by 32.5% to 14,722. With the exception of 2012, a consistent increase was observed across all years of birth, with the largest increase seen in the first year and estimates appearing to stabilise within 5 years of the initial report. Excluding 2012, the average percentage increase (shown by the black line in Figure 5.4) was 21.7% between the initial and next report, and 32.3% by 5 years later.

Table 5.4 - Method 2: Percentage increase in the reported number of live births each year, by years since initial estimate

		Years since initial estimate					
		0	1	2	3	4	5
Calendar year of birth	2010	0.0%	19.6%	24.6%	28.9%	31.3%	32.3%
	2011	0.0%	20.8%	28.9%	32.5%	33.1%	
	2012	0.0%	34.3%	39.7%	25.5%	54.3%	
	2013	0.0%	23.9%	26.1%	32.1%		
	2014	0.0%	18.1%	25.4%			
	2015	0.0%	25.8%				
	2016	0.0%					
Average		0.0%	23.8%	28.9%	29.7%	39.6%	32.3%
Average (excluding 2012)		0.0%	21.7%	26.3%	31.2%	32.2%	32.3%

Figure 5.4 - Method 2: Percentage increase in the reported number of live births each year, by years since initial estimate
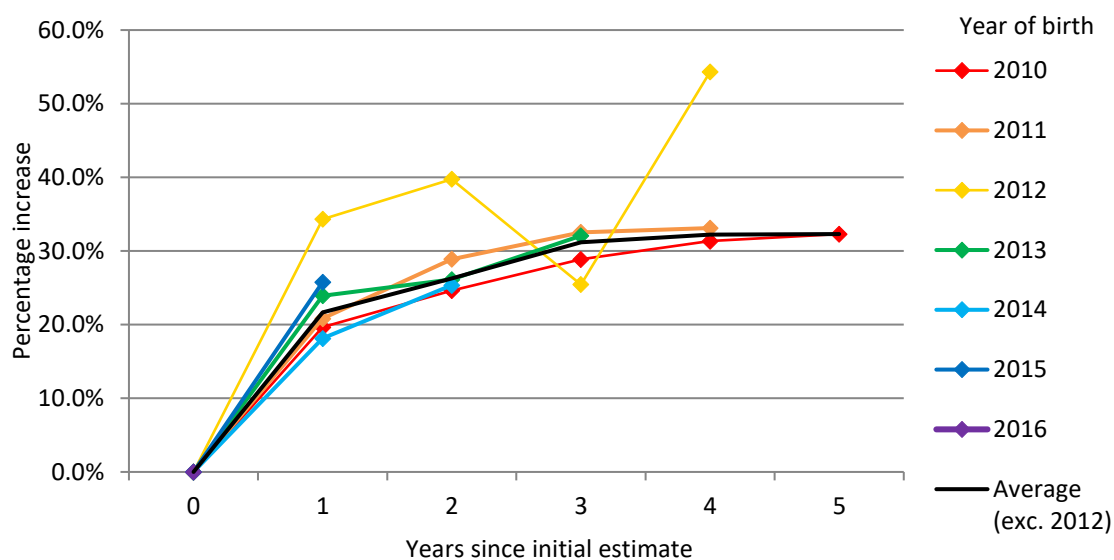
Based on these percentages, reported numbers were adjusted to obtain a more accurate estimate of the true number of births each year, as shown in Table 5.5. For example, the most recent estimate of the number of births in 2014 came from the 2016 report, 2 years after the initial estimate. The adjusted estimate was therefore calculated by first dividing by 1.263 (as the average percentage increase between the initial estimate and the one 2 years later was 26.3%), and then multiplying by 1.323 (as the average percentage increase from the initial report to the one 5 years later, our best estimate of the true figure, was 32.3%). This gave a final adjusted estimate of 17,533 births in 2014, an increase of 795 over the most recently reported estimate for this year. The estimate for 2010 was unchanged, as this year already had 5 years of published follow-up data on late registrations and so no further adjustment could be made.

Estimates of the antenatal seroprevalence in the district were taken from ANCHSS, and were adjusted to reflect incident HIV in pregnancy subsequent to the HIV test for the ANCHSS surveillance. The rate of incident HIV during pregnancy in the sub-district was 0.0008654 (95% CI 0.0006538, 0.0011154) per person-week. MONARCH data suggested that the mean gestational age was 19.5 weeks at first ANC visit, relevant for the 2010 to 2014 ANCHSS surveys, and 37.5 weeks at one randomly selected visit per woman, relevant for the 2015 survey year. The risk of a woman acquiring HIV between sampling and delivery was therefore 0.0008654*(40-19.5)=1.77% for 2010 to 2014, and 0.0008654*(40-37.5)=0.22% for 2015. Using the confidence interval for the estimated rate of incident HIV during pregnancy, lower and upper bounds of the risk of HIV between sampling and delivery were calculated (using same methods as above) to be 1.34% and 2.29% for 2010 to 2014, and 0.16% and 0.28% for 2015.

Table 5.5 - Method 2: Adjusted estimates of the number of live births each year in uMkhanyakude based on data from SSA

Calendar year of birth	Initially reported estimate	Most recently reported estimate	Years between initial and most recent estimates	Adjustment of most recent estimate	Adjusted estimate	Increase from most recently reported to adjusted estimate
Source of data	SSA	SSA		Table 5.4		
Method of calculation	A	B		C	D=B*C	D-B
2010	10,592	14,013	5	-	14,013	0
2011	11,107	14,785	4	x (1.323/1.322)	14,796	11
2012	11,722	18,088	4	x (1.323/1.322)	18,102	14
2013	13,231	17,474	3	x (1.323/1.312)	17,621	147
2014	13,353	16,738	2	x (1.323/1.263)	17,533	795
2015	12,992	16,340	1	x (1.323/1.217)	17,763	1,423
2016	13,858	13,858	0	x 1.323	18,334	4,476

The original estimates of antenatal HIV prevalence from the ANCHSS and the adjusted estimates (including those based on the above lower and upper bounds) are shown in Table 5.6. For example, in 2013 the ANCHSS estimated that the antenatal HIV seroprevalence was 44.1%. Among the 55.9% who did not test positive at sampling, 1.77% would be expected to acquire HIV before delivery, corresponding to 0.99% of the total number of pregnant women. The total proportion of women who were living with HIV by delivery was therefore estimated to be 45.09%, with lower and upper bounds of 44.85% and 45.38% respectively.

Table 5.6 - Method 2: Original and adjusted estimates of antenatal seroprevalence in uMkhanyakude, 2010-2016

Calendar year of birth	Proportion of women positive in survey	Additional proportion of women becoming positive by delivery	Proportion of women positive by delivery		
			Estimate	Lower bound	Upper bound
Source of data	ANCHSS				
Method of calculation	A	B=(1-A)*(Risk of HIV per week*median number of weeks)	A+B		
2010	41.9%	(1-41.9%)*1.77%=1.03%	41.9%+1.03%=42.93%	42.68%	43.23%
2011	41.1%	(1-41.1%)*1.77%=1.04%	41.1%+1.04%=42.14%	41.89%	42.45%
2012	35.2%	(1-35.2%)*1.77%=1.15%	35.2%+1.15%=36.35%	36.07%	36.68%
2013	44.1%	(1-44.1%)*1.77%=0.99%	44.1%+0.99%=45.09%	44.85%	45.38%
2014	39.9%	(1-39.9%)*1.77%=1.06%	39.9%+1.06%=40.96%	40.71%	41.28%
2015	46.3%	(1-46.3%)*0.22%=0.0001%	46.3%+0.0001%=46.30%	46.39%	46.45%
2016	46.3%	(1-46.3%)*0.22%=0.0001%	46.3%+0.0001%=46.30%	46.39%	46.45%

Note: Due to a lag in reporting, no estimate of the antenatal seroprevalence was available for 2016, so the estimate for 2015 was used

The final estimates of the proportion of infants tested for method 2 are shown in Table 5.7. The final estimate of the number of HIV-exposed infants born in the sub-district by calendar year is shown in column E, calculated as the product of the number of live births, the antenatal seroprevalence in uMkhanyakude, the estimated population of the sub-district as a proportion of uMkhanyakude, and the amount of follow-up time included each year. The overall proportion increased from 86% under the old guidelines to 89% after the introduction of birth testing in April 2015 (p<0.001), and the proportion by 7 weeks of age from 45% to 57% (p<0.001).

Testing coverage based on unadjusted estimates in the calculation of the number of HIV-exposed infants and of duplicated PCR test data are shown in Table 5.8. Across the whole time period the estimated number of HIV-exposed infants decreased from 17,570 to 16,137 and the number tested increased from 15,234 to 17,622, resulting in an increase in the estimate of the proportion ever tested over the whole time period from 87% to 109%.

Method 3 - NHLS-DHIS

The number of live births to women with HIV reported to the DHIS are shown in column B of Table 5.9, and the adjusted numbers allowing for births outside healthcare facilities (3% of all births) are shown in column D. The proportion of infants ever tested is shown in the final column, and was 115% between April 2014 and March 2015, increasing to 153% in the period to December 2016. The estimated number of live births to women with HIV was much lower with Method 3 compared to Method 2 (for example, for 2016, 1,578 compared to 3,097). The estimated seroprevalence (calculated based on the reported total number of live births, shown in column A, and the reported number of live births to women with HIV) was relatively similar to that estimated in Method 2.

Testing coverage based on unadjusted estimates in the calculation of the number of HIV-exposed infants and of duplicated PCR test data is shown in Table 5.10; compared to the estimates based on the adjusted data, the proportion ever tested across the whole time period increased from 138% to 167% and by 7 weeks of age from 86% to 91%.

Table 5.7 - Method 2: Estimates of the proportion of HIV-exposed infants who received a PCR test based on adjusted NHLS-SSA/ANCHSS data, overall, by year of birth and by guideline time period

Guideline time period	Calendar year of birth	Number of live births in uMkhanyakude in whole year	Antenatal seroprevalence	Population of Hlabisa sub-district as proportion of uMkhanyakude	Time (years)	Number of HIV-exposed infants born in Hlabisa sub-district	Number of infants tested	Number of infants tested by 7 weeks of age	Proportion of infants ever tested	Proportion of infants tested by 7 weeks of age
	Source of data	Table 5.5	Table 5.6	[132, 209]			NHLS	NHLS		
	Method of calculation	A	B	C	D	E=A*B*C*D	F	G	(F/E)*100	(G/E)*100
Prior to introduction of birth testing	June - December 2010	14,013	42.93%	36.48%	7/12	1,280	1,044	517	82%	40%
	2011	14,796	42.14%	36.48%	1	2,275	2,084	1,119	92%	49%
	2012	18,102	36.35%	36.48%	1	2,400	2,234	1,102	93%	46%
	2013	17,621	45.09%	36.48%	1	2,898	2,173	1,081	75%	37%
	2014	17,533	40.96%	36.48%	1	2,620	2,373	1,329	91%	51%
	January - March 2015	17,763	46.30%	36.48%	3/12	750	582	320	78%	43%
	Total					12,223	10,490	5,468	86%	45%
After introduction of birth testing	April - December 2015	17,763	46.30%	36.48%	9/12	2,250	2,221	1,463	99%	65%
	2016	18,334	46.30%*	36.48%	1	3,097	2,523	1,605	81%	52%
	Total					5,347	4,744	3,068	89%	57%
TOTAL						17,570	15,234	8,536	87%	49%

*No estimate of the antenatal seroprevalence was available for 2016, so the estimate for 2015 was used

Table 5.8 - Method 2: Estimates of the proportion of HIV-exposed infants who received a PCR test based on unadjusted NHLS-SSA/ANCHSS data, overall, by year of birth and by guideline time period

Guideline time period	Calendar year of birth	Number of live births in uMkhanyakude in whole year, as reported by SSA	Antenatal seroprevalence, as reported by ANCHSS	Population of Hlabisa sub-district as proportion of uMkhanyakude	Time (years)	Unadjusted number of HIV-exposed infants born in Hlabisa sub-district	Number of infants tested, prior to deduplication	Number of infants tested by 7 weeks of age, prior to deduplication	Proportion of infants ever tested (unadjusted)	Proportion of infants tested by 7 weeks of age (unadjusted)
	Source of data	SSA	ANCHSS	[132, 209]			NHLS	NHLS		
	Method of calculation	A	B	C	D	E=A*B*C*D	F	G	(F/E)*100	(G/E)*100
Prior to introduction of birth testing	June - December 2010	14,013	41.90%	36.48%	7/12	1,249	1,181	522	95%	42%
	2011	14,785	41.10%	36.48%	1	2,217	2,408	1,123	109%	51%
	2012	18,088	35.20%	36.48%	1	2,323	2,569	1,109	111%	48%
	2013	17,474	44.10%	36.48%	1	2,811	2,450	1,084	87%	39%
	2014	16,738	39.90%	36.48%	1	2,436	2,667	1,332	109%	55%
	January - March 2015	16,340	46.30%	36.48%	3/12	690	650	321	94%	47%
	Total					11,726	11,925	5,491	102%	47%
After introduction of birth testing	April - December 2015	16,340	46.30%	36.48%	9/12	2,070	2,678	1,524	129%	74%
	2016	13,858	46.30%*	36.48%	1	2,341	3,019	1,658	129%	71%
	Total					4,411	5,697	3,182	129%	72%
TOTAL						16,137	17,622	8,673	109%	54%

*No estimate of the antenatal seroprevalence was available for 2016, so the estimate for 2015 was used

Table 5.9 - Method 3: Estimates of the proportion of HIV-exposed infants who received a PCR test based on adjusted NHLS-DHIS data, overall, by year of birth and by guideline time period

Guideline time period	Calendar year of birth	Total number of live births in facilities	Number of live births to women with HIV in facilities	Estimated seroprevalence	Number of live births to women with HIV, adjusted for births outside facilities	Number of infants tested	Number of infants tested by 7 weeks of age	Proportion of infants ever tested	Proportion of infants tested by 7 weeks of age
	Source of data	DHIS	DHIS	DHIS		NHLS	NHLS		
	Method of calculation	A	B	C=B/A	D=B/0.97	E	F	(E/D)*100	(F/D)*100
Prior to introduction of birth testing	April - December 2014	3,898	1,395	35.79%	1,438	1,737	1,026	121%	71%
	January - March 2015	1,428	561	39.29%	578	582	320	101%	55%
	Total				2,016	2,319	1,346	115%	67%
After introduction of birth testing	April - December 2015	3,778	1,475	39.04%	1,521	2,221	1,463	146%	96%
	2016	4,703	1,531	32.55%	1,578	2,523	1,605	160%	102%
	Total				3,099	4,744	3,068	153%	99%
TOTAL					5,115	7,063	8,536	138%	86%

Table 5.10 - Method 3: Estimates of the proportion of HIV-exposed infants who received a PCR test based on unadjusted NHLS-DHIS data, overall, by year of birth and by guideline time period

Guideline time period	Calendar year of birth	Total number of live births in facilities	Number of live births to women with HIV in facilities	Estimated seroprevalence	Number of infants tested, prior to deduplication	Number of infants tested by 7 weeks of age, prior to deduplication	Proportion of infants ever tested (unadjusted)	Proportion of infants tested by 7 weeks of age (unadjusted)
	Source of data	DHIS	DHIS	DHIS	NHLS	NHLS		
	Method of calculation	A	B	C=B/A	E	F	(E/B)*100	(F/B)*100
Prior to introduction of birth testing	April - December 2014	3,898	1,395	35.79%	1,952	1,028	140%	74%
	January - March 2015	1,428	561	39.29%	650	321	116%	57%
	Total				2,602	1,349	132%	69%
After introduction of birth testing	April - December 2015	3,778	1,475	39.04%	2,678	1,524	182%	103%
	2016	4,703	1,531	32.55%	3,019	1,658	197%	108%
	Total				5,697	3,182	190%	106%
TOTAL					8,299	4,531	167%	91%

<u>Method 4 - Road-to-health booklets (MONARCH)</u>

The number of women and infants in the MONARCH trial is shown in Figure 5.5. A total of 3,171 women attended a study clinic for antenatal care and were therefore included in the trial, of whom 2,551 (80%) had data captured from their maternal case record including data on HIV testing; 1,156 (45%) were known to be living with HIV by the time of their delivery, 1,384 (54%) were HIV negative, and the remaining 11 (<1%) had unknown HIV status. Among women with maternal case record data and across all HIV exposure groups, 1,827 (72%) had available delivery outcome data (corresponding to 1,827/3,171 (58%) of all women included in the trial), to whom 1,848 infants were born. Of the 837 HIV-exposed infants, 819 (98%) ever had their booklet photographed, of whom 405 (49%) had their booklet seen once, 314 (38%) twice, and 100 (12%) at all three time points. The number whose booklet was photographed at delivery was 623 (76%), at 3-6 days 420 (51%), and at 6 weeks 290 (35%).

Estimates of PCR testing coverage are shown by step of the stepped wedge trial design, in Table 5.11 and by year of birth in Table 5.12. Of 819 HIV-exposed infants, 380 (46%) had a record of a PCR test in their booklet. Of these, the date of the PCR test was available for 22 infants and the result was available for 4 infants. An additional 21 infants born to women who did not have HIV also had a recorded PCR test. The proportion of infants with a recorded PCR test declined markedly over the course of the trial, from 88% in step 0 (the baseline data collection phase), to 26% in step 7. Among those born in 2015 the proportion tested was 72%, compared to 34% among those born in 2016. There was some evidence that the proportion with a PCR test was higher among those whose booklet was photographed at 6 weeks compared to those whose booklet was not (147/290 (51%) vs. 233/529 (44%), p=0.068).

In order to assess the completeness of reporting of PCR test data in Road-to-health booklets, the number of PCR tests reported to have been conducted at the 7 demographic surveillance area clinics in MONARCH was compared to the number in the NHLS database in Table 5.13, by step. As only infants born to 58% of the women included in MONARCH had maternal case record and Road-to-health booklet data available, it might be expected that the number of tests recorded in the MONARCH dataset would be only 58% of the number reported in the NHLS data, however the number actually reported was much less than this. Looking primarily at steps 0, 1, 2 (where data from all clinics in the demographic surveillance area were available), the number of tests in MONARCH relative to NHLS decreased over time.

Figure 5.5 - Method 4: Flowchart of women and infants included in the MONARCH trial

3,171 women included in trial

2,551 (80%) women with a maternal case record

620 (20%) women with no maternal case record

1,156 (45%) women with HIV

1,384 (54%) women HIV-negative

11 (<1%) women with unknown HIV status

827 (72%) with at least one linked infant record

994 (72%) with at least one linked infant record

6 (54%) with at least one linked infant record

total = 1,827 women

837 infants

1,005 infants

6 infants

total = 1,848 infants

819 (98%) infants whose road-to-health booklet was ever photographed

980 (98%) infants whose road-to-health booklet was ever photographed

6 (100%) infants whose road-to-health booklet was ever photographed

total = 1,805 infants

Table 5.11 - Method 4: Estimates of the proportion of HIV-exposed infants who received a PCR test based on data from Road-to-health booklets (MONARCH) by step in the stepped wedge design of the trial

Step	Infant's date of birth	Estimate of PCR coverage
0	15th July 2015 - 28th September 2015	129/147 (88%)
1	29th September 2015 - 23rd November 2015	55/86 (64%)
2	24th November 2015 - 26th January 2016	32/83 (39%)
3	27th January 2016 - 16th March 2016	42/86 (49%)
4	17th March 2016 - 17th May 2016	35/95 (37%)
5	18th May 2016 - 18th July 2016	35/98 (36%)
6	19th July 2016 - 22nd September 2016	23/114 (20%)
7	23rd September 2016 - 30th January 2017	27/104 (26%)
	Unknown date of birth	2/6 (33%)
	Total	380/819 (46%)

Table 5.12 - Method 4: Estimates of the proportion of HIV-exposed infants who received a PCR test based on data from Road-to-health booklets (MONARCH), overall and by year of birth

Guideline time period	Calendar year of birth	Number of HIV-exposed infants	Number with a recorded PCR test	Proportion of infants tested by 7 weeks of age
	Source of data	MONARCH	MONARCH	
	Method of calculation	A	B	(B/A)*100
After introduction of birth testing	July - December 2015	265	192	72%
	2016	548	186	34%
TOTAL		813	378	46%

Note: Excludes 6 HIV-exposed infants with unknown date of birth

Table 5.13 - Method 4: Comparison of number of infant HIV PCR tests conducted at 7 demographic surveillance area clinics reported in the MONARCH and NHLS datasets

Step	Infant's date of birth	Number of days	Road-to-health booklets (MONARCH)	NHLS (to 7 weeks of age only)	MONARCH /NHLS
0	15th July 2015 - 28th September 2015	76	129	206	63%
1	29th September 2015 - 23th November 2015	56	55	136	40%
2	24th November 2015 - 26th January 2016	65	32	161	20%
3	27th January 2016 - 16th March 2016	50	42	72	58%
4	17th March 2016 - 17th May 2016	62	35	67	52%
5	18th May 2016 - 18th July 2016	62	35	98	36%
6	19th July 2016 - 22nd September 2016	66	23	177	13%
7	23rd September 2016 - 30th January 2017	130	27	239	11%

<u>Comparison of the four methods:</u>

A comparison of all estimates from all four methods is shown in Table 5.14 and graphically in Figure 5.6. The estimates of the proportion of infants ever tested ranged considerably across the first three methods, with those from method 1 being the lowest, at between 29% and 48% by year, followed by method 2, at between 75% and 99%, and method 3 the highest, at between 101% and 160%. A similar pattern was observed in the proportion tested by 7 weeks of age, with the corresponding estimate made using method 4 in between the other three.

Figure 5.6 - Comparison of estimates of the proportion of HIV-exposed infants who received a PCR test using methods 1-4, both (a) ever and (b) by 7 weeks of age
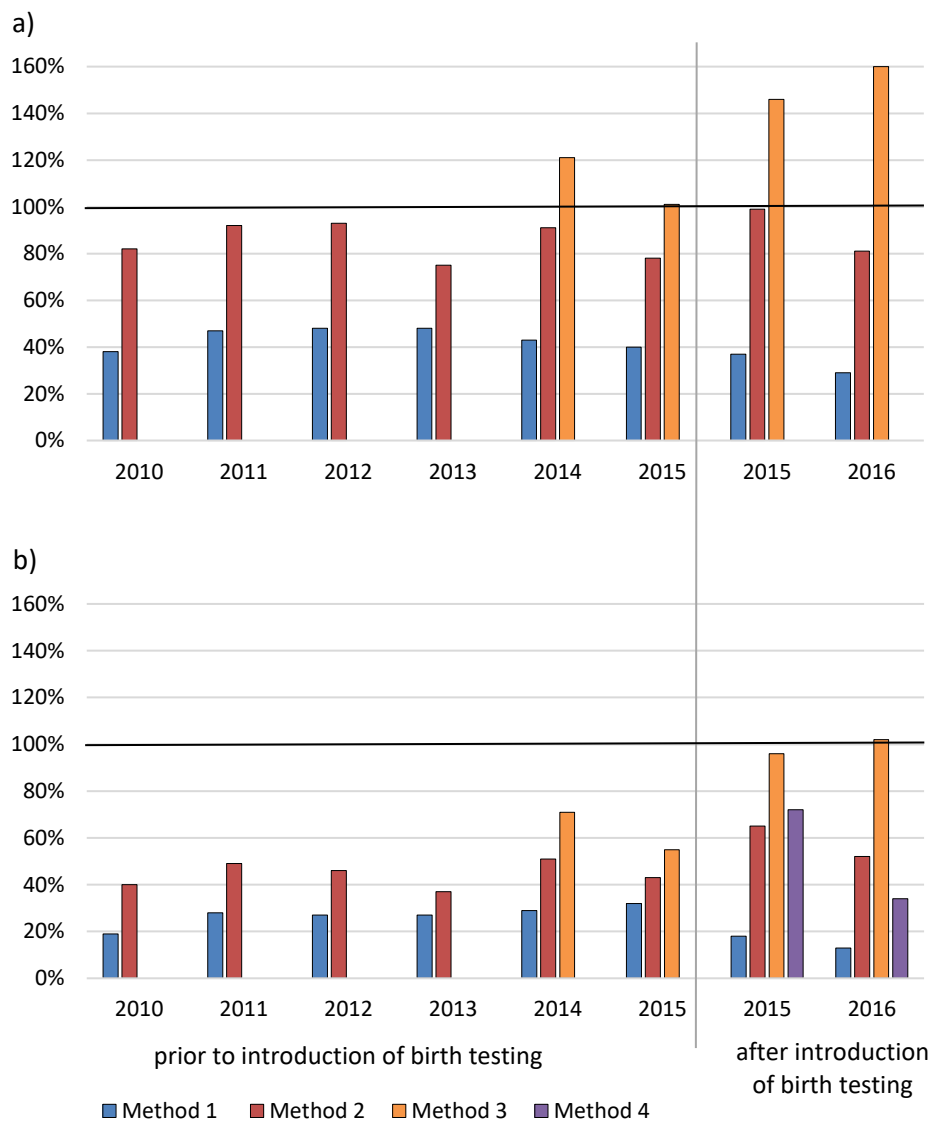
Table 5.14 - Comparison of estimates of the proportion of HIV-exposed infants who received a PCR test using methods 1-4, both ever and by 7 weeks of age

Guideline time period	Calendar year of birth	Proportion of infants ever tested according to method:			Proportion of infants tested by 7 weeks of age according to method:			
		1	2	3	1	2	3	4
Prior to introduction of birth testing	June - December 2010	38%	82%	-	19%	40%	-	-
	2011	47%	92%	-	28%	49%	-	-
	2012	48%	93%	-	27%	46%	-	-
	2013	48%	75%	-	27%	37%	-	-
	2014	43%	91%	121%	29%	51%	71%	-
	January - March 2015	40%	78%	101%	32%	43%	55%	-
After introduction of birth testing	April - December 2015	37%	99%	146%	18%	65%	96%	72%
	2016	29%	81%	160%	13%	52%	102%	34%

Notes: Estimates only available from April 2014 onwards for method 3 and from July 2015 onwards for method 4. Method 1=NHLS-AHRI surveillance, Method 2=NHLS-SSA/ANCHSS, Method 3=NHLS-DHIS, Method 4=Road-to-health booklets (MONARCH).

5.5. Discussion

In this chapter, I used several sources of data and four different methods to estimate the proportion of HIV-exposed infants in the sub-district who received an HIV PCR test. The estimates of the proportion ever tested ranged considerably across the first three methods, with those from NHLS-AHRI surveillance being the lowest, at between 29% and 48% by year, followed by NHLS-SSA/ANCHSS, at between 75% and 99%, and NHLS-DHIS the highest, at between 101% and 160%. There was no clear trend by year with any of these methods. Data available in MONARCH could only be used to estimate the proportion of infants tested by their 6 week post-natal visit, with estimates for this method somewhere in the middle of those calculated for this same outcome with the other methods.

A summary of the limitations of each method, and the possible resulting direction of biases, is shown in Table 5.15.

Methods 1 to 3 all used NHLS data to estimate the number of infants who received a PCR test, and therefore all relied on the 'within NHLS linkage' conducted in stage 1 of Chapter 4 (where repeat tests on the same infant were identified) to estimate the numerator of the number of infants tested. Failure to link repeat tests from the same infant due to, for example, the use of the mother's name on the NHLS form rather than the infant's for early tests, may have resulted in an overestimation of the number of infants ever tested and thus an overestimation of the true testing coverage. Conversely, although less likely given the limitations of the data, it is possible that tests on different infants may have been erroneously linked, resulting in the underestimation of the true coverage.

Table 5.15 - Comparison of the 4 methods for estimating HIV PCR testing coverage in HIV-exposed infants

	Method	Data source	Issue	Type of error on numerator/ denominator	Likely direction of bias on estimate of coverage	Adjustment made
Numerator (PCR tests)	1, 2, 3	NHLS PCR dataset	NHLS data limitations may have led to failure to identify repeat tests on the same infant	Misclassification	Overestimation	
			Tests from different infants may be linked	Misclassification	Underestimation	
			Data from one clinic only available between January 2011 and March 2016	Missing data	Underestimation	
			Not possible to estimate the number of infants who received care at private healthcare facilities	Misclassification	Underestimation	
			Inclusion of PCR tests conducted on infants born to women who acquired HIV in postpartum period (who are not included in the denominator)	Misclassification	Overestimation	
	1	NHLS-ACDIS linkage	Failure to link infants to their PCR tests	Misclassification	Underestimation	
	4	Recording of PCR tests in Road-to-health booklets	Poor completion of PCR data in Road-to-health booklet	Misclassification	Underestimation	
			Not all infants had their Road-to-health booklet photographed at their 6 week visit	Selection bias	Underestimation	
Denominator (HIV-exposed infants)	1	Women in AHRI surveillance known to have HIV	Relies on AHRI's ACDIS-ACCDB linkage to include women on ART in the denominator	Misclassification	Either	
			Only women on ART or willing to participate in serosurvey can be included	Selection bias	Either	
			Small number of women included	Reliability	Either	
	2	SSA/ANCHSS	Seroprevalence estimate based on small number of survey participants in each district	Reliability	Either	
			Women acquiring HIV after sampling for survey not included	Misclassification	Overestimation	Estimate of incident HIV during pregnancy used to adjust seroprevalence

2	SSA/ANCHSS	Underestimation of number births due to late registration	Misclassification	Overestimation	Change in reported numbers over time used to adjust estimates for more recent years
		Numerator not directly linked to denominator	Validity	Either	
3	DHIS	Known underreporting of births in DHIS data	Misclassification	Overestimation	
		Infants born outside of healthcare facility not included	Misclassification	Overestimation	Numbers adjusted for infants born at home
		Numerator not directly linked to denominator (overall number of PCR tests conducted not available in DHIS)	Validity	Either	
4	Infants in MONARCH with Road-to-health booklet available	Not all infants had Road-to-health booklet photographed (possibly as a result of mother not giving consent)	Selection bias	Either	

In addition to the methodological difficulties of linkage within the NHLS dataset, errors in the extraction of data from the NHLS data warehouse meant that data from one clinic were only available from 1st January 2011 to 20th March 2016. Some infants may have been tested at this clinic outside of this time period and would therefore not be counted as tested in this analysis, resulting in the underestimation of coverage. An additional limitation to the use of the NHLS dataset is that it only includes tests conducted at public healthcare facilities, while each of methods 1, 2 and 3 would also count infants who received care in private facilities in the denominator, leading to possible underestimation of coverage.

A further limitation to the use of NHLS PCR data is that although the denominator used in each analysis included only in utero HIV-exposed infants, it is possible that some infants born to women who acquired HIV during the postpartum period would also have received a PCR test and therefore have been counted in the numerator, leading to the over inflation of testing coverage for methods 2 and 3. However, since women are not routinely screened for HIV in the postpartum period, it is unlikely that many of these small number of infants would have been identified, and thus it is unlikely that this would have a major impact on the results.

Method 1 estimated coverage for infants born to women identified as having HIV either through the AHRI serosurvey or through linkage to ACCDB. Estimates of testing coverage based on this method were the lowest of all 4 methods, suggesting the underestimation of the true proportion of infants tested. As with the 'within NHLS linkage', some links between infants in ACDIS and their NHLS PCR test data in stage 3 of Chapter 4 are likely to have been missed, resulting in misclassification of some infants who actually did receive a PCR test. Fewer identifying variables were available for this stage of the linkage than the others, increasing the chance of missing links.

The number of HIV-exposed infants found to be living in the demographic surveillance area in method 1 was less than expected; approximately 40% of all those living within the entire Hlabisa sub-district live within the AHRI demographic surveillance area [132, 133], however only 6% (965/15,234) of infants tested in the entire sub-district were identified here. There are several possible reasons for this. Firstly, only infants born to women with HIV who tested positive in the serosurvey or were linked to ACCDB could be identified. Between 2003 and 2012, participation in the serosurvey was estimated at between 26% and 46% in any given year (though 77% participated at least once in the first 9 years after becoming eligible), with having HIV known to be associated with non-participation [130]. In addition, missed matches in AHRI's linkage between ACDIS and ACCDB may have meant that not all women on ART were identified; only 72% of mothers with HIV in ACDIS were also found in ACCDB, despite ART coverage among pregnant women in the area being estimated at 93% [138]. Secondly, even if a woman was identified as having HIV, the date of HIV acquisition used in analysis (the earliest of the date of

first positive serosurvey test or date of ART start) might have been much later than the true date of infection, so although all infants born to a woman would have been identified in ACDIS data collection, the HIV exposure status of some may have been incorrect.

It is possible that the women (and their infants) who were included in the denominator for the analysis for method 1 were not representative of all women with HIV in the sub-district, resulting in selection bias. Maternal characteristics including not knowing HIV status prior to delivery and poor adherence to treatment, which would have been associated with reduced likelihood of identification through the serosurvey and linkage to ACCDB, are known to be associated with poor uptake of EID services [220]. Infants born to women included in this analysis may therefore be more likely to receive a PCR test, resulting in the overestimation of the true testing coverage. Conversely, since until 2013 only those pregnant women with CD4 counts below 350 cells/mm^3 were eligible for treatment, the women included in this analysis would be sicker than average; this may have resulted in either the underestimation or overestimation of coverage in earlier years. Finally, a relatively small number of women and their infants were included in the denominator for method 1 compared to the other methods, which may affect the reliability of the estimates.

Method 2 used estimates of the number of live births from SSA to calculate the denominator. The number of live births each year reported by SSA was adjusted to account for births which were registered late and therefore not included, although it is possible that the true number of infants born was still higher than that estimated here, which would result in the overestimation of PCR testing coverage. Barriers to registration include some parents having to travel long distances to Department of Home Affairs offices [221], although recently steps have been taken to increase timeliness of registration, for example, increasing the number of on-site registration facilities in hospitals [193]. A survey in another rural area of South Africa showed that although completeness of birth registration has increased rapidly over time from only 6.2% in 1992 to 70.0% in 2010, it had still only reached 90.5% by 2014 [222]. An underestimation of the number of births of even by 10% would still result in a significant change to coverage estimates here.

Antenatal seroprevalence estimates in method 2 came from ANCHSS. Although participation in the survey was offered to all women attending care at the chosen sampled clinics, informed consent was required; this may have introduced selection bias, as women who knew they had HIV may have been less likely to agree to participate, resulting in underestimation of the true prevalence (and thus overestimation of testing coverage). In addition, some women may have acquired HIV after participating in the survey, resulting in further underestimation of the prevalence. Although it was possible to account for some of this in analysis using the estimated rate of incident HIV during pregnancy in the sub-district, infants exposed to HIV through breastfeeding as a result of women newly acquiring HIV in the postpartum period were not

accounted for. The estimates of seroprevalence varied substantially by year (for example from 41% to 35% to 44% in 2011, 2012 and 2013 respectively), resulting in high variation in the corresponding estimates of testing coverage. It seems unlikely that this reflects true variation in the prevalence over time, and may be unreliability caused by the relatively small number of clinics sampled, and therefore the number of women sampled, when stratifying by district.

Estimates of coverage from DHIS (method 3) exceeded 100% for all calendar years, averaging 115% under the before the introduction of birth testing and 153% after the introduction. This suggests a huge underestimation of the number of HIV-exposed infants born within the sub-district, with reported estimates approximately half of those calculated using data from SSA/ANCHSS in method 2, even after adjustment for births outside of healthcare facilities. Both the number of live births and the antenatal seroprevalence (calculated as the number of live births to women with HIV divided by the total number of live births) from DHIS were lower than estimates from SSA/ANCHSS, suggesting not all births were recorded, with more births to women with HIV missing. Despite data quality assessment processes, unreliable DHIS data collection has been reported in other studies [123] and hypothesised causes of error include poorly designed data collection tools, incorrect transcription of values, and staff not knowing the definition and/or existence of all data items to be collected [223]. It is also possible that some records were not actually sent from clinics to DHIS for processing. A further weakness of the DHIS data is that while some information on PCR testing at specific ages is collected, the total number of PCR tests conducted is not, so this source could not be used to calculate the numerator of the coverage.

An additional weakness for methods 2 and 3 is the use of data sources for the calculation of the numerator and denominator which were not directly linked, rather than determining the achievement of the outcome of receipt of a PCR test at an individual-level, which is generally preferable, giving internal consistency [92].

Estimation of testing coverage using data from the MONARCH trial (method 4) relied upon complete recording of PCR test data in an infant's Road-to-health booklets for the numerator. A previous cross-sectional study in Pretoria, South Africa in 2012 observed poor recording of HIV-related data, with, for example, 24% of known HIV-exposed infants having no record of maternal HIV status. It is known that parents often request HIV-related data be omitted or written in code to protect confidentiality if the booklet has to be taken into school or might be seen by other family members [224]. In my analysis, even among infants who had receipt of a PCR test recorded in their booklet, the actual date of the test was only recorded 6% of the time. Similarly, although not all results would have been returned to the clinic by time of photographing the booklet, only 1% of tested infants had a test result recorded. It is also possible that some data were recorded in the booklet, but were not correctly entered into the MONARCH database. This

poor reporting may explain the lower estimates of coverage obtained here compared to some of the other methods, although it should be noted that coverage was slightly higher (51% vs. 44%) among those whose booklet was seen at the 6 week visit compared to those only seen at delivery and/or the 3-6 day visit.

Although all women attending for antenatal care within the demographic surveillance area were eligible for inclusion in MONARCH, only 58% of those included in the trial had data on both maternal HIV status and infant follow-up to enable their inclusion in this analysis. Although photographs of the Road-to-health booklet were not available for all infants, they were missing for a similar proportion of those in the HIV-exposed and HIV-unexposed groups, suggesting that the missingness was not related to HIV status and therefore not the likelihood of receiving a PCR test. In addition to missing Road-to-health booklet data, 20% of women had no maternal case record data available to determine their HIV status. Possible reasons for this include women giving birth at a clinic outside of the study area, or refusal.

The proportion of infants tested in MONARCH was highest in the early part of the trial, at 88% in the baseline data collection phase, decreasing rapidly to less than 30% by the end of the trial. The reasons for this are unclear. The same decrease was not observed when looking at the number of tests conducted in the demographic surveillance area clinics according to the NHLS dataset, suggesting that it may at least partly be a result of worsening recording of data in the Road-to-health booklet as the trial progressed, with staff practice possibly reverting to standard pre-trial levels at the end of the follow-up period. It is also possible that the observed decline was a result of the study intervention, with staff focusing on maternal HIV and viral load testing (which were the trial's primary outcomes).

The strengths and weaknesses discussed above highlight the difficulties in estimating PCR testing coverage, with high variation between the estimates from each of the methods, meaning that the true coverage remains unclear. Given the known issues with DHIS data collection, and that the resulting estimates made with method 3 were above 100% for each calendar year, this method may be considered the most unreliable. Similarly, it seems unlikely that estimates from method 4, which was based on Road-to-health booklet data collected in the MONARCH trial, were correct, given the lower than expected number of PCR tests reported compared to the NHLS dataset. The true testing coverage may therefore lie between the estimates calculated using methods 1 and 2, but it is difficult to assess which of these is the most accurate; method 2 may have overestimated coverage because of difficulties obtaining an accurate estimate of number of births, and it is likely that missed links between ACDIS and NHLS in method 1 would have resulted in underestimation.

Other studies from South Africa have estimated testing coverage using methods similar to method 2 here. One estimated national EID coverage (defined as the proportion tested by 2 months of age), but the authors were unable to deduplicate repeat tests on the same infant in the NHLS database [123]. The estimated coverage increased steadily over time using this method, from 52% in 2010 to 87% in 2014. These estimates were higher than those calculated by 7 weeks of age using method 2 in my analysis, but were closer to the estimates based on unadjusted antenatal prevalence and live births and PCR data prior to deduplication.

The estimate of testing coverage to 2 months of age for South Africa from UNAIDS was 79% by 2016 [22], markedly higher than the estimates made here. UNAIDS estimates however varied year-on-year and were greater than 100% in some years, for example from 114% in 2015 to 79% in 2016 to 101% in 2017, further highlighting the methodological difficulties of measuring this outcome [225].

The methods used here were based on data from a variety of both routinely collected data sources as well as data collected specifically for research, although none were without limitations and the true testing coverage remains uncertain. It is necessary to improve monitoring of this important target to evaluate how well guidelines are being implemented at the local level, in particular in areas of high maternal HIV prevalence such as this. Few studies consider the limitations of sources of routine data to the level of detail discussed here. One previous study compared estimates made using two methods (comparable to methods 2 and 3 used here) and explored the effect of different definitions of coverage based on the DHIS data [123]. However, this earlier study did not consider all limitations, for example, the authors did not deduplicate infants with repeat PCR tests within the NHLS dataset. In broader public health terms, it is critical to ensure that infants at risk of HIV are appropriately followed and tested after birth, and to eliminate the risk of new infections in infants going undiagnosed, with such infections being associated with high mortality [226, 227].

5.6. Key findings

The key findings from this chapter are:

- All routine data sources have limitations, which can have a significant impact on the results of analyses based on them
- Even with access to a variety of different data sources and methods, it is difficult to accurately assess coverage of PCR testing of HIV-exposed infants, and thus whether the guidelines are being followed
- The true testing coverage remains unclear, but is unlikely to be close to 100%.

Chapter 6. Factors associated with HIV PCR testing

6.1. Introduction

In Chapter 5, I showed that estimates of PCR testing coverage varied considerably by data source, and it was hard to infer the true level of coverage. Notwithstanding this limitation, it may still be possible to use the datasets which contained individual-level patient data to explore whether any groups of infants were more or less likely to be tested. Thus, in this chapter, I use data from two of the methods from Chapter 5 to explore associations between maternal and infant characteristics and the receipt of PCR testing.

6.2. Objectives

The objective for this chapter is to:

- Explore associations between key maternal, pregnancy and delivery related characteristics and coverage of infant HIV PCR testing.

6.3. Methods

Of the four methods used to estimate PCR testing coverage in Chapter 5, two, methods 1 (based on data from NHLS-AHRI surveillance) and 4 (based on data collected for the MONARCH trial), used individual-level data with a linked numerator and denominator, and had available information characterising the infants and their mothers. These methods could therefore be used to investigate factors associated with PCR testing, the binary outcome of interest. For method 1 (NHLS-AHRI), the outcome was categorised as ever received a PCR test vs. never received a PCR test, and for method 4 (MONARCH), as received a PCR test by the 6 week postnatal visit vs. did not receive a PCR by the 6 week postnatal visit (as no data were collected after this timepoint).

The variables available within each dataset are shown in Table 6.1. For method 1, characteristics were derived from ACDIS. Maternal education was based on the highest education level reported by the time of delivery. Maternal marital status, employment status, household asset index and beliefs about ART (where participants were asked whether ART can help people improve their health) are also collected as part of the annual survey, with the value reported at the survey visit closest to delivery (within a window of ±1 year) used in the analysis. Household asset index is calculated by AHRI based on the ownership of various household assets and type of access to amenities such as drinking water, and is then scaled and categorised into quintiles (across the whole population). The number of previous live-born children was categorised at 0, 1, 2 or ≥3, given the small number of women with more than 3 children. Women were

considered to be on ART if they had ever initiated ART as indicated by having been linked to ACCDB, with CD4 and viral load (categorised as ≤/>400 copies/mL) at delivery (allowing a window of ±6 months) determined through ACCDB for those who were on ART. Time since ART initiation was categorised as more than or less than 4 weeks. This was not available as a continuous variable, as the actual date of initiation was not available for women who transferred to a clinic in the sub-district already on ART (although no women transferred in during the last 4 weeks of pregnancy, enabling categorisation as less than or more than 4 weeks of ART). Time since maternal HIV diagnosis was calculated with the date of HIV diagnosis defined as the earliest of the date first reported to be on ART in ACCDB or the date they first tested positive in the HIV serosurvey. Data on hospitalisations during pregnancy came from linkage between ACDIS and Hospital Information System (HIS) which had already been conducted by AHRI, with hospitalisations that began on the infant's date of birth, which were assumed to be for delivery, excluded.

For method 4, based on data collected for the MONARCH trial, the number of previous live-born children born to each mother (subsequently categorised as 0, 1, 2, ≥3), maternal ART history, whether they were diagnosed with HIV before or during pregnancy, the number of antenatal clinic visits and CD4 and viral load (categorised as ≤/>400 copies/mL) measurements (again, with the most recent within a window of ±6 months used) were collected as recorded within the maternal case record. Data on gestational age at delivery (subsequently categorised as </≥ 37 weeks), infant's date of birth and delivery location (categorised as public clinic, public hospital, home or other) were collected from Road-to-health booklets. Data on maternal education and employment came from questions asked at the 6 week post-natal interview, as did beliefs about ART improving health (with participants asked whether the risk of falling ill was reduced with ARV treatment).

No clinic-level characteristics (for example, the size of the clinic at which the infant was born) could be included in either analysis for the following reasons: for the infants included in the analysis based on method 1 who did not receive a PCR test, no data were available on which facility they were born in or received care at. For method 4, although this information was available, only approximately 30% of infants were born in a clinic.

Infant and maternal characteristics from each method were described using medians and interquartile ranges (IQR) for continuous covariates, and frequencies and percentages for categorical covariates. Univariable associations between each factor and PCR testing were explored using a chi-squared test (or Fisher's exact test if there were fewer than five infants in any group) for categorical covariates, and Wilcoxon's rank-sum test for continuous covariates, excluding those with missing values.

Table 6.1 - Covariates potentially associated with PCR testing and their availability in the NHLS-AHRI surveillance and MONARCH datasets

Category	Covariate	Variable type	Method 1: NHLS-AHRI surveillance	Method 4: MONARCH
Calendar time	Infant's date of birth	Continuous	✓	✓
Maternal sociodemographic	Age at time of delivery	Continuous	✓	✓
	Education	Categorical	✓	✓
	Employment	Categorical	✓	✓
	Marital status	Categorical	✓	✓
	Number of previous live-born children	Categorical	✓	✓
	Household asset index	Categorical	✓	✗
Maternal HIV-related	Time since HIV diagnosis	Continuous (M1) Categorical (M4)	✓	✓
	Any ART use prior to delivery	Binary	✓	✓
	CD4 count at delivery	Continuous	✓	✓
	Virally suppressed at delivery	Binary	✓	✓
	Beliefs about ART and health	Categorical	✓	✓
Pregnancy/ delivery	Number of antenatal clinic visits	Continuous	✗	✓
	Hospital admissions during pregnancy	Binary	✓	✗
	Gestational age <37 weeks at delivery	Binary	✗	✓
	Delivery location	Categorical	✓	✓

Variable type listed is as included in analysis, rather than as collected

To further investigate the association between each factor and PCR testing, a multivariable analysis was conducted for each dataset. Before building the models, associations between the factors in each of the datasets were explored to investigate whether any were highly collinear. The inclusion of highly collinear covariates in a multivariable model can lead to instability, whereby small changes in the data result in large changes to the estimated effect sizes. In addition, the standard errors for the coefficients of the affected covariates may be large, resulting in wide confidence intervals [228]. In this situation, it may be difficult to identify the covariates with the strongest effect. In order to determine collinearity, Cramer's V test was used, with a threshold of >0.5 considered evidence of strong association [229]. For the purpose of this assessment, continuous covariates were categorised into quartiles. If two collinear covariates were identified, the one with the strongest univariable association with PCR testing was kept for analysis.

Different methods for dealing with missing data in the multivariable model were considered. The missing indicator method involves including an extra dummy variable in the model for each factor, to indicate whether the value is missing for that covariate, allowing all individuals to be included in the analysis. However, this can introduce bias into the estimates of non-missing levels of covariates, even if the data are missing completely at random [230]. Further, this requires that all continuous covariates with missing data be categorised, which may not always

be appropriate. A complete case analysis, that is, including only individuals with complete data on both the outcome and all covariates in the model, was therefore considered. Covariates with data missing for ≥10% of individuals were excluded, so that a reasonable number of individuals were included in analysis.

A complete case analysis gives unbiased estimates of the logistic regression coefficients in the situation where the probability of being a complete case does not depend on the outcome (after accounting for the other covariates) [231, 232]. In each of the analyses here, it was not explicitly reported if the infant did not have a PCR test, rather it was assumed that no test was carried out if a PCR test was not linked to an infant in case of the NHLS-AHRI analysis or recorded in the infant's Road-to-health booklet in the case of the MONARCH analysis. Consequently, there were no missing data in the outcome variable, and so being a complete case was defined as having available data on all the covariates. Several statistical tests were used to assess whether the probability of being a complete case was associated with having a PCR test. Firstly, a chi-squared test was used, with no evidence of an association for either method; in the NHLS-AHRI dataset there were 2,037 (90%) complete cases, consisting of 91% of those with a PCR test vs. 90% of those without (p=0.479), and in the MONARCH dataset there were 781 (95%) complete cases, consisting of 96% of those tested compared to 95% of those not (p=0.833). Subsequently, a logistic regression model was fitted, with being a complete case as the outcome and whether the infant received a PCR test as a predictor, adjusted for the other covariates. Adding all covariates into this model together would have resulted in only individuals with complete data being included, and thus the perfect prediction of the outcome. Only complete covariates were therefore included; there was no evidence of association between PCR testing and being a complete case. Lastly, a model with PCR testing as the outcome, the complete covariates, and an indicator for the missingness of each of the incomplete covariates in turn was run to explore the association between PCR testing and the missingness of each covariate; there was again no evidence of an association observed. A complete case analysis was therefore deemed unbiased.

In each dataset, some women had more than one infant. In the NHLS-AHRI dataset (where women may have had more than one child either as a result of twins or triplets, or because of more than one pregnancy over the time period covered) there were 2,254 infants born to 1,918 unique mothers, and in the MONARCH dataset (where, given the shorter time period covered, no women had more than one pregnancy) there were 819 infants with 810 unique mothers. This violates the assumption of independent observations which is required for logistic regression [233]. Two alternative methods of modelling the data were therefore considered. A conditional model (such as a generalized linear mixed model), estimates subject-specific effects, while a marginal model (such as a generalized estimating equations approach, GEE), estimates the average effect of a covariate across the population [234]. Given the interest in identifying

characteristics associated with low PCR testing uptake across the population, a GEE model was used for this analysis. An exchangeable covariance structure was used, making the assumption of equal correlation between all of the infants born to each mother.

The strategy used to build a multivariable model depends on the purpose of model [235]; in this analysis the purpose was to describe and explore relationships between all of the available covariates and PCR testing, rather than to build a predictive model or to investigate the effect of one particular covariate of interest (adjusting for the others) on testing. In order to interpret the estimated effect of each covariate when controlling for the others, it is important to understand the relationships between covariates. To obtain an unbiased estimate of the causal effect of an exposure of interest on the outcome, one should adjust for all confounders (that is, covariates associated with the exposure and that are independently associated with the outcome) but not mediators (that is, covariates which lie on the causal pathway between the exposure and the outcome) [236]. One approach is to use a conceptual hierarchical framework to describe the assumed relationships between the covariates to be included in the model, as a guide to model building. This involves allocating the covariates to different hierarchical levels, depending on their mechanism of effect on the outcome [237], with the following four levels used in this analysis:

1. Pregnancy/delivery related factors (the most proximal to the outcome)
2. Maternal HIV-related factors
3. Maternal sociodemographic factors
4. Calendar time (the most distal to the outcome)

The covariates were allocated to each of these levels as shown in Table 6.1.

The assumed relationships between covariates in each of these levels and the outcome of PCR testing is shown in Figure 6.1. It was assumed that all factors have an effect on the outcome, some of which is direct and some of which is mediated through the factors in the more proximal hierarchical levels. Calendar time (included in the analysis using the infant's date of birth) has some direct effect on PCR testing (arrow (a)), for example as a result of changes to the national PCR testing guidelines over time, but also an effect mediated through each of the other more proximal factors (arrows (b)-(d)) as each of these may change over time, for example household asset index has improved across the sub-district over time. It was assumed that the sociodemographic factors have some direct effect on the outcome (arrow (e)), that some of their effect is mediated through the other more proximal factors, both HIV-related (arrow (g)) and pregnancy/delivery related (arrow (f)) (for example, maternal education level may influence beliefs about the relationship between ART and health outcomes), but that none is mediated through calendar time (as a change in a sociodemographic factor cannot cause a change in

calendar time). Similarly, maternal HIV-related factors were assumed have some direct effect (arrow (h)), as well as some mediated through the pregnancy/delivery related factors (arrow (i)). Finally, it was assumed that all of the effect of the pregnancy/delivery related factors on the outcome was direct (arrow (j)), with none mediated through either calendar time, the HIV-related or the sociodemographic factors.

Figure 6.1 - Conceptual hierarchical framework for PCR testing
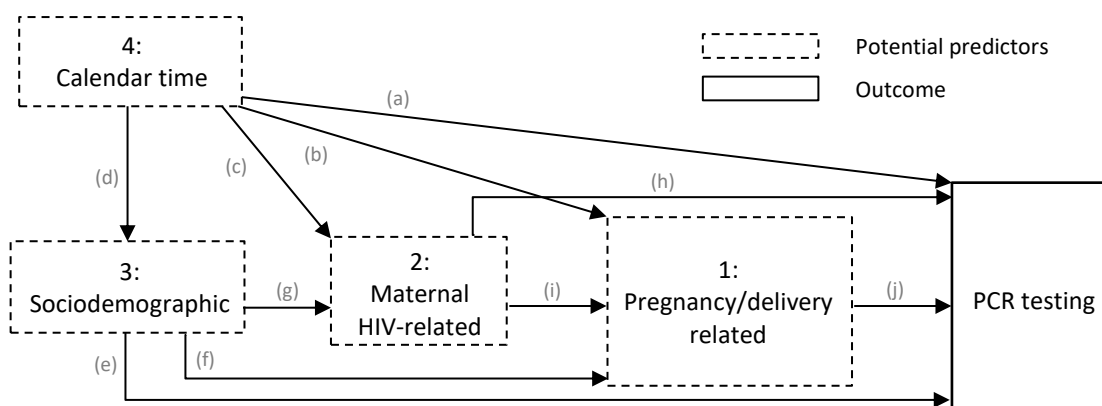
When building the multivariable model, the effect of each covariate was estimated adjusting for factors within the same level of the hierarchy and for those in more distal levels, but not those in the more proximal ones, that is, not adjusting for those with at least some mediating effect on the outcome. Four multivariable models were therefore built for each analysis. The first just included terms for the infant's date of birth. The second, used to assess the effect of each of the sociodemographic factors, included just the sociodemographic factors, as well as calendar time. The third, used to assess the effect of each of the maternal HIV-related factors, adjusted for calendar time, the sociodemographic factors and the other HIV-related factors. Finally the fourth, used to assess the effect of the pregnancy/delivery related covariates, adjusted for all other factors.

All the covariates within each level of the hierarchy were included in the relevant final models, rather than carrying out variable selection, for example based on a threshold of having a p-value less than 0.05. This was deemed appropriate, given the reasonable sample size, the small number of covariates considered, and that the proportion of infants tested was not close to 0 or 1 [238].

In order to determine the most appropriate functional form of the continuous covariates included in the model, multivariable models including a linear term and using restricted cubic splines (with 3 knots at the 10th, 50th and 90th percentiles [239]) were compared using quasi-likelihood under the independence model criterion (QIC), with the form with the lower QIC value chosen [240, 241]. For continuous covariates represented with spline terms in the final model,

odds ratios were presented at key values. Interactions between the infant's date of birth and each of the other covariates included in the multivariable models were explored to investigate changes in their effects over time.

6.4. Results

A summary of the characteristics of women and infants identified with each method is shown in Table 6.2. Characteristics were generally similar between the two groups, with three exceptions. The proportion of women who had initiated ART by delivery was lower among those identified through NHLS-AHRI surveillance compared to those included in MONARCH (62% vs. 97%), although the proportion on ART in the NHLS-AHRI dataset increased (to 75%) when restricting to the same calendar time period as the MONARCH trial. The median CD4 count at delivery was lower among women in MONARCH (460 vs. 520 cells/mm^3), although over half of women in each dataset had no measurement available. A lower proportion of women in MONARCH reported agreement that 'ART improves health' than in the NHLS-AHRI data (82% vs. 98%), although the wording of the questions asked differed.

Results from the univariable analyses using data from NHLS-AHRI surveillance are shown in Table 6.3. Infants born in earlier calendar years were more likely to be tested (the median date of birth for those tested was April 2013 vs. November 2013 for those not tested, p=0.001), as were those born to older mothers (median age 28.9 for those tested vs. 28.1 years for those not, p=0.021) and to infants of mothers who had started ART prior to delivery (45% vs. 40% for those whose mothers had no started ART, p=0.033).

Results from univariable analyses using data from MONARCH are shown in Table 6.4. Infants born in earlier calendar years were more likely to have received a PCR test (p<0.001). Infants born to mothers with two previous live-born children were more likely to be tested than those born to mothers with none, one or three or more (p=0.029). Infants with mothers who had initiated ART less than 4 weeks at delivery were more likely to be tested than those who had started ART before pregnancy or who had started during pregnancy but more than 4 weeks ago (81% vs. 40% vs. 48% respectively, p<0.001). Those born at full term were more likely to be tested than those born pre-term (47% vs. 35%, p=0.062). Finally, those born in a clinic were more likely to be tested than those born in a hospital or at home (65% vs. 38% and 55% respectively, p<0.001).

Table 6.2 - Characteristics of those identified through NHLS-AHRI surveillance and MONARCH

		NHLS-AHRI surveillance (N=2,254)	MONARCH (N=819)
		n (column %) or median (IQR)	
Calendar time	Infant's date of birth (0% missing, 1% missing)		
	Continuous	Jul 13 (Dec 11, Dec 14)	Mar 16 (Nov 15, Jul 16)
Maternal sociodemographic	Age at time of delivery, years (0% missing, 2% missing)	28.4 (24.0, 32.8)	28.3 (23.8, 32.6)
	Education (13% missing, 32% missing)		
	Still in school	44 (2%)	2 (<0.5%)
	Did not complete secondary education	869 (44%)	265 (47%)
	Completed secondary education	671 (34%)	252 (45%)
	Tertiary education	372 (19%)	44 (8%)
	Employed (23% missing, 30% missing)		
	Still in school	44 (2%)	2 (<0.5%)
	Part time employment	120 (7%)	105 (18%)
	Full time employment	265 (15%)	
	No employment	1,351 (76%)	475 (82%)
	Marital status (1% missing, 10% missing)		
	Married	119 (5%)	39 (5%)
	Single	-	667 (90%)
	Stable relationship	-	31 (4%)
	Other	-	1 (<0.5%)
	Not married	127 (6%)	-
	Widowed	12 (1%)	-
	Separated/divorced	2 (<0.5%)	-
	Informal union	1,964 (88%)	-
	Number of previous live-born children (2% missing, 1% missing)		
	0	420 (19%)	171 (21%)
	1	648 (29%)	294 (36%)
	2	538 (24%)	209 (26%)
	≥3	601 (27%)	137 (17%)

141

Maternal sociodemographic	Household asset index (5% missing, -)		
	1st quintile	224 (10%)	-
	2nd quintile	345 (16%)	-
	3rd quintile	493 (23%)	-
	4th quintile	611 (28%)	-
	5th quintile	474 (22%)	-
Maternal HIV-related	Time since HIV diagnosis by delivery (0% missing, 1% missing)		
	Diagnosed before pregnancy	-	463 (57%)
	Diagnosed during pregnancy	-	346 (43%)
	Continuous, years	2.5 (0.6, 5.6)	-
	Any ART use prior to delivery (0% missing, 0% missing)	1403 (62%)	794 (97%)
	Duration of ART use by delivery among those on ART, years (1% missing, 0% missing)		
	Started prior to pregnancy	1,349 (96%)	336 (42%)
	Started during pregnancy, had >4 weeks ART	1,349 (96%)	401 (51%)
	Started during pregnancy, had <4 weeks ART	49 (4%)	57 (7%)
	CD4 count at delivery among those on ART, cells/mm^3 (77% missing, 54% missing)	520 (359, 675)	460 (346, 649)
	Viral load at delivery among those on ART (80% missing, 74% missing)		
	≤400 copies/mL	240 (84%)	183 (89%)
	>400 copies/mL	47 (16%)	23 (11%)
	Believes ART improves health (21% missing, 64% missing)		
	Agree	1,743 (98%)	240 (82%)
	Disagree	12 (1%)	36 (12%)
	Don't know	30 (2%)	18 (6%)
Pregnancy/ Delivery	Number of antenatal clinic visits (-, 0% missing)	-	6 (5, 8)
	Hospital admissions during pregnancy, excluding delivery (0% missing, -)	23 (1%)	-
	Gestational age at delivery (-, 10% missing)		
	Pre-term (<37 weeks)	-	69 (9%)
	Full term (≥37 weeks)	-	666 (91%)

Pregnancy/ Delivery	Delivery location (4% missing, 0% missing)		
	Public clinic	601 (28%)	251 (31%)
	Public hospital	1451 (68%)	546 (67%)
	Home	72 (3%)	20 (2%)
	Private facility	22 (1%)	0
	Other	0	2 (<0.5%)

'-' indicates category not included within this dataset

Table 6.3 - Associations between covariates and PCR testing, using data from NHLS-AHRI surveillance

		Total (N=2,254)	Ever received a PCR test (N=965, 43%)	Never received a PCR test (N=1,289, 57%)	p-value
			n or n (row %) or median (IQR)		
Calendar time	Infant's date of birth				
	Continuous	Jul 13 (Dec 11, Dec 14)	Apr 13 (Dec 11, Sep 14)	Nov 13 (Jan 12, Dec 15)	0.001
	Age at time of delivery				
	Continuous, years	28.4 (24.0, 32.8)	28.9 (24.4, 33.1)	28.1 (23.7, 32.7)	0.021
	Education*				
	Still in school	44	23 (52%)	21 (48%)	
	Did not complete secondary education	869	372 (43%)	497 (57%)	
	Completed secondary education	671	275 (41%)	396 (59%)	0.388
	Tertiary education	372	166 (45%)	206 (55%)	
	Unknown	298	129 (43%)	169 (57%)	
	Employed*				
	Still in school	44	23 (52%)	21 (48%)	
	Part time employment	120	60 (50%)	60 (50%)	
Maternal sociodemographic	Full time employment	265	108 (41%)	157 (59%)	0.159
	No employment	1,351	563 (42%)	788 (58%)	
	Unknown	474	211 (45%)	263 (55%)	
	Marital status				
	In a relationship	2,083	899 (43%)	1,184 (57%)	
	Not in a relationship	141	53 (38%)	88 (62%)	0.196
	Unknown	30	13 (43%)	17 (57%)	
	Number of previous live-born children				
	0	432	175 (41%)	257 (59%)	
	1	643	271 (42%)	372 (58%)	
	2	533	236 (44%)	297 (56%)	0.418
	≥3	599	271 (45%)	328 (55%)	
	Unknown	47	12 (26%)	35 (74%)	

Table 6.3 continued on the next page…

Maternal sociodemographic	Household asset quintile				
	1st quintile	224	111 (50%)	113 (50%)	
	2nd quintile	345	153 (44%)	192 (56%)	
	3rd quintile	493	194 (39%)	299 (61%)	0.205
	4th quintile	611	256 (42%)	355 (58%)	
	5th quintile	474	206 (43%)	268 (57%)	
	Unknown	107	45 (42%)	62 (58%)	
Maternal HIV-related	Time since HIV diagnosis by delivery				
	Continuous, years	2.5 (0.6, 5.6)	2.7 (0.7, 5.8)	2.4 (0.6, 5.5)	0.505
	Any ART use prior to delivery				
	Yes	1,403	625 (45%)	778 (55%)	0.033
	No	851	340 (40%)	511 (60%)	
	Duration of ART use by delivery among those on ART				
	Had >4 weeks ART	1,349	602 (45%)	747 (55%)	
	Had <4 weeks ART	49	21 (43%)	28 (57%)	0.807
	Unknown	5	2 (40%)	3 (60%)	
	CD4 count at delivery among those on ART*				
	Continuous, cells/mm^3	520 (359, 675)	557 (375, 700)	504 (345, 660)	0.352
	Unknown	1,041	464 (45%)	577 (55%)	
	Viral load at delivery among those on ART*				
	≤400 copies/mL	244	110 (45%)	134 (55%)	
	>400 copies/mL	43	19 (44%)	26 (56%)	0.913
	Unknown	1,116	496 (44%)	620 (56%)	
	Believes ART improves health*				
	Agree	1,743	755 (43%)	988 (57%)	
	Disagree	12	5 (42%)	7 (58%)	0.762
	Don't know	30	11 (37%)	19 (63%)	
	Unknown	469	194 (41%)	275 (59%)	

…Table 6.3 continued

	Admitted to hospital during pregnancy, excluding delivery				
	Yes	23	9 (39%)	14 (61%)	0.720
	No	2,231	956 (43%)	1,296 (57%)	
Pregnancy/ delivery	Delivery location				
	Public clinic	601	260 (43%)	341 (57%)	
	Public hospital	1,451	631 (43%)	820 (57%)	0.413
	Home	72	28 (39%)	44 (61%)	
	Private facility	22	6 (27%)	16 (73%)	
	Unknown	108	40 (37%)	68 (63%)	

*Excluded from the multivariable analysis as a result of the high proportion (≥10%) of missing data

Table 6.4 - Associations between covariates and PCR testing, using data from MONARCH

		Total (N=819)	Ever received a PCR test (N=380, 46%)	Never received a PCR test (N=439, 54%)	p-value
			n or n (row %) or median (IQR)		
Calendar time	Infant's date of birth				
	Continuous	Mar 16 (Nov 15, Jul 16)	Dec 15 (Sep 15, Apr 16)	Jun 16 (Feb 16, Aug 16)	<0.001
	Unknown	6	2 (33%)	4 (67%)	
Maternal sociodemographic	Age at time of delivery				
	Continuous, years	28.3 (23.8, 32.6)	29.0 (24.0, 33.0)	27.8 (23.7, 32.3)	0.134
	Unknown	20	13 (65%)	7 (35%)	
	Education*				
	Still in school	2	1 (50%)	1 (50%)	
	Did not complete secondary education	265	124 (47%)	141 (53%)	
	Completed secondary education	250	130 (52%)	120 (48%)	0.739
	Tertiary education	44	22 (50%)	22 (50%)	
	Unknown	258	103 (40%)	155 (60%)	
	Employed*				
	Still in school	2	1 (50%)	1 (50%)	
	Employed	105	55 (52%)	50 (48%)	0.725
	Not employed	475	227 (48%)	248 (52%)	
	Unknown	237	97 (41%)	140 (59%)	
	Marital status*				
	Single	667	312 (47%)	355 (53%)	
	Married	39	13 (33%)	26 (67%)	
	Stable relationship	31	17 (55%)	14 (45%)	0.191
	Other	1	1 (100%)	0	
	Unknown	81	37 (46%)	44 (54%)	

	Number of previous live-born children				
Maternal sociodemographic	0	171	73 (43%)	98 (57%)	
	1	294	133 (45%)	161 (55%)	0.029
	2	209	115 (55%)	94 (45%)	
	≥3	137	56 (41%)	81 (59%)	
	Unknown	8	3 (38%)	5 (63%)	
	Timing of HIV diagnosis				
	Diagnosed before pregnancy	463	208 (45%)	255 (55%)	0.205
	Diagnosed during pregnancy	346	171 (49%)	175 (51%)	
	Unknown	10	1 (10%)	9 (90%)	
	Any ART use prior to delivery				
	Yes	794	372 (47%)	422 (53%)	0.143
	No	25	8 (32%)	17 (68%)	
	Duration of ART use by delivery among those on ART				
	Started prior to pregnancy	336	134 (40%)	202 (60%)	
	Started during pregnancy, had >4 weeks ART	401	192 (48%)	209 (52%)	<0.001
Maternal HIV-related	Started during pregnancy, had <4 weeks ART	57	46 (81%)	11 (19%)	
	CD4 count at delivery among those on ART*				
	Continuous, cells/mm^3	460 (346, 649)	453 (347, 614)	467 (341, 678)	0.386
	Unknown	439	213 (49%)	226 (51%)	
	Viral load at delivery among those on ART*				
	≤400 copies/mL	183	68 (37%)	115 (63%)	0.313
	>400 copies/mL	23	7 (30%)	16 (70%)	
	Unknown	588	297 (51%)	291 (49%)	
	Believes ART improves health				
	Agree	240	119 (50%)	121 (50%)	
	Disagree	36	21 (58%)	15 (42%)	0.618
	Don't know	18	9 (50%)	9 (50%)	
	Unknown	525	231 (44%)	294 (56%)	

147

	Number of antenatal clinic visits				
	Continuous	6 (5, 8)	6 (5, 8)	7 (5, 9)	0.885
	Gestational age at delivery*				
	Pre-term (<37 weeks)	69	24 (35%)	45 (65%)	0.062
Pregnancy/ delivery	Full term (≥37 weeks)	666	310 (47%)	356 (53%)	
	Unknown	84	46 (55%)	38 (45%)	
	Delivery location				
	Public clinic	251	164 (65%)	87 (35%)	
	Public hospital	546	205 (38%)	341 (62%)	<0.001
	Home	20	11 (55%)	9 (45%)	
	Other	2	0	2 (100%)	

*Excluded from the multivariable analysis as a result of the high proportion (≥10%) of missing data

The results from the univariable and multivariable logistic regression analyses based on NHLS-AHRI surveillance data are shown in Table 6.5. The following covariates had more than 10% of values missing, and were therefore excluded from the multivariable analysis: education and employment status, CD4 and viral load at delivery, and beliefs about ART improving health. Of the total 2,254 infants in the dataset, 2,037 (90%) had complete data on the remaining covariates and were included in the analysis (compared to only 194 (9%) who would have been included if only those with complete data on all covariates had been included). The covariates for ART use prior to delivery and the duration of ART use were combined to make one covariate with the following categories: no prior ART use, <4 weeks since ART initiation or ≥4 weeks since ART initiation. No covariates were found to be collinear, with a Cramer's V statistic less than 0.5 for each pair. There was no evidence of a non-linear effect of either maternal age (QIC 2,825.9 for the model with a linear term vs. 2,827.6 with restricted cubic splines) or time since HIV diagnosis (QIC 2,818.6 vs. 2,819.1) so both were included as linear terms, but there was of infant's date of birth (QIC 3,070.1 vs. 3,051.5), which was therefore included using restricted cubic spline terms.

The odds of being tested increased with more recent date of infant's birth up to 2013 (adjusted odds ratio (aOR) for an infant born on 1st July 2013 1.40 (95% CI 1.05, 1.89) vs. 1st July 2010), before subsequently decreasing after 2013 (aOR for an infant born on 1st July 2016 0.58 (95% CI 0.43, 0.77) vs. 1st July 2010) (p<0.001). After adjustment for infant's date of birth and other sociodemographic factors, there was weak evidence that testing coverage increased with maternal age (aOR 1.02 per year older, 95% CI 1.00, 1.03, p=0.096). After adjustment for sociodemographic and HIV-related factors, infants whose mother had initiated ART were more likely to receive a PCR test than those not (aOR for ≥4 weeks ART 1.29 (95% CI 1.07, 1.57), and for <4 weeks ART 1.22 (0.66, 2.25), vs. no ART, p=0.033). There was no evidence of a statistically significant effect of any of the other covariates, and there was no evidence of an interaction between infant's date of birth and any of the other covariates.

Table 6.5 - Hierarchical multivariable logistic regression of factors associated with PCR testing, based on data from NHLS-AHRI surveillance

	Univariable			Multivariable		
	OR	95% CI	p	aOR	95% CI	p
CALENDAR TIME*						
Infant's date of birth						
1st July 2010				1.00	-	
1st July 2011				1.20	1.06, 1.36	
1st July 2012				1.39	1.10, 1.76	
1st July 2013		-		1.40	1.05, 1.89	<0.001
1st July 2014				1.16	0.87, 1.54	
1st July 2015				0.84	0.64, 1.09	
1st July 2016				0.58	0.43, 0.77	
MATERNAL SOCIODEMOGRAPHIC**						
Mother's age, per year increase	1.01	1.00, 1.03	0.065	1.02	1.00, 1.03	0.096
Marital status						
In a relationship	1.00	-	0.191	1.00	-	0.955
Not in a relationship	0.79	0.55, 1.12		0.99	0.67, 1.45	
Number of previous live born children						
0	1.00	-		1.00	-	
1	1.06	0.83, 1.36	0.432	0.96	0.73, 1.27	0.987
2	1.16	0.89, 1.50		0.98	0.73, 1.32	
≥3	1.21	0.94, 1.56		1.00	0.72, 1.39	
Household asset quintile						
1st quintile	1.51	1.10, 2.08		1.43	1.03, 2.00	
2nd quintile	1.22	0.93, 1.62		1.20	0.90, 1.60	
3rd quintile	1.00	-	0.136	1.00	-	0.304
4th quintile	1.11	0.87, 1.41		1.13	0.88, 1.45	
5th quintile	1.17	0.90, 1.51		1.13	0.87, 1.47	
MATERNAL HIV-RELATED***						
Time since mother's HIV diagnosis at delivery, per year increase	1.01	0.98, 1.03	0.678	1.00	0.97, 1.03	0.898
Prior maternal ART use at time of delivery						
None	1.00	-		1.00	-	
<4 weeks	1.16	0.65, 2.08	0.088	1.22	0.66, 2.25	0.033
≥4 weeks	1.22	1.02, 1.45		1.29	1.07, 1.57	
DELIVERY/PREGNANCY****						
Hospital admissions during pregnancy (excluding for delivery)						
Yes	0.87	0.37, 2.03	0.751	0.74	0.30, 1.82	0.513
No	1.00	-		1.00	-	
Delivery location						
Public clinic	1.00	-		1.00	-	
Public hospital	1.01	0.83, 1.22	0.429	1.04	0.85, 1.28	0.280
Home	0.83	0.50, 1.37		0.82	0.49, 1.37	
Private facility	0.49	0.19, 1.28		0.45	0.17, 1.18	

OR: odds ratio; aOR: adjusted odds ratio.
Infant's date of birth was included in the model using restricted cubic spline terms; odds ratios are presented for the 1st July each year, relative to the 1st July 2010.
*Univariable model the same as the multivariable model; no adjustment made for other factors as calendar time is the most distal factor
**Multivariable estimates adjusted for infant's date of birth, and other sociodemographic factors
***Multivariable estimates adjusted for infant's date of birth, sociodemographic factors, and other HIV-related factors
****Multivariable estimates adjusted for infant's date of birth, sociodemographic factors, HIV-related factors, and other delivery/pregnancy related factors

Table 6.6 - Hierarchical multivariable logistic regression of factors associated with PCR testing, based on data from MONARCH

	Univariable			Multivariable		
	OR	95% CI	p	aOR	95% CI	p
CALENDAR TIME*						
Infant's date of birth						
15th July 2015				1.00	-	
15th December 2015		-		0.16	0.10, 0.25	<0.001
15th July 2016				0.05	0.03, 0.08	
15th December 2016				0.04	0.02, 0.08	
MATERNAL SOCIODEMOGRAPHIC**						
Mother's age, per year increase	0.99	0.95, 1.05	0.337	1.01	0.98, 1.05	0.438
Number of previous live-born children						
0	1.00	-		1.00	-	
1	1.08	0.74, 1.58	0.040	1.21	0.77, 1.90	0.003
2	1.63	1.08, 2.45		2.09	1.23, 3.52	
≥3	0.96	0.61, 1.53		0.94	0.50, 1.74	
MATERNAL HIV-RELATED**						
Timing of maternal HIV diagnosis						
Prior to pregnancy	0.84	0.63, 1.11	0.208	0.76	0.54, 1.09	0.133
During pregnancy	1.00	-		1.00	-	
Prior maternal ART use at time of delivery						
None	1.00	-		1.00	-	
<4 weeks	8.89	3.06, 25.83	<0.001	2.32	0.67, 8.03	0.186
≥4 weeks	1.69	0.72, 3.98		1.15	0.42, 3.15	
DELIVERY/PREGNANCY**						
Number of antenatal care visits, per unit increase	0.99	0.94, 1.05	0.828	1.00	0.92, 1.08	0.945
Effect of delivery location						
Public clinic	1.00	-				
Public hospital	0.32	0.23, 0.44	<0.001	p-value for interaction: <0.001		
Other	0.53	0.22, 1.27				
Effect of delivery location on 15th July 2015						
Public clinic				1.00	-	
Public hospital		-		20.82	6.86, 63.21	<0.001
Other				0.84	0.67, 10.64	
Effect of delivery location on 15th December 2015						
Public clinic				1.00	-	
Public hospital		-		0.43	0.27, 0.70	<0.001
Other				0.32	0.08, 1.22	
Effect of delivery location on 15th July 2016						
Public clinic				1.00	-	
Public hospital		-		0.06	0.04, 0.10	<0.001
Other				0.40	0.12, 1.34	
Effect of delivery location on 15th December 2016						
Public clinic				1.00	-	
Public hospital		-		0.14	0.04, 0.44	<0.001
Other				1.29	0.06, 25.64	

Infant's date of birth was included in the model using restricted cubic spline terms; odds ratios are presented at dates chosen to represent shape of association, relative to 15th July 2015.

*Univariable model the same as the multivariable model; no adjustment made for other factors as calendar time is the most distal factor

**Multivariable estimates adjusted for infant's date of birth, and other sociodemographic factors

***Multivariable estimates adjusted for infant's date of birth, sociodemographic factors, and other HIV-related factors

****Multivariable estimates adjusted for infant's date of birth, sociodemographic factors, HIV-related factors, and other delivery/pregnancy related factors

The results from the univariable and multivariable logistic regression analyses based on data from the MONARCH trial are shown in Table 6.6 and Figure 6.2. The following covariates had 10% or more of values missing and were therefore excluded from the multivariable analysis: education, employment, marital status, CD4 and viral load at delivery, beliefs about ART improving health, and gestational age at delivery. Of the total 819 infants in the dataset, 781 (95%) had complete data on the remaining covariates and were included in the analysis (compared to only 32 (4%) who would have been included if all covariates had been used). No covariates were found to be collinear, with a Cramer's V statistic less than 0.5 for each pair. There was no evidence of a non-linear effect of either maternal age (QIC 930.9 for the model with a linear term vs. 932.8 with restricted cubic splines) or number of antenatal care visits (QIC 854.4 vs. 855.6) so both were included as linear terms, but there was of infant's date of birth (QIC 976.8 vs. 959.9), which was therefore included using restricted cubic spline terms.

The odds of receiving a PCR test decreased for those born in more recent years; relative to those born on 15th July 2015, the aORs (95% CI) for those born on 15th December 2015, 15th July 2016 and 15th December 2016 were 0.16 (0.10, 0.25), 0.05 (0.03, 0.08) and 0.04 (0.02, 0.08) respectively. Adjusting for calendar time and maternal age, infants born to women with two previous children were most likely to be tested (aOR 2.09, 95% CI 1.23, 3.52). There was evidence of a change in the effect of delivery location over time (p<0.001), adjusting for all other factors (Figure 6.2). Earlier on in the trial infants born in a hospital were more likely to receive a PCR test compared to those born in a clinic (aOR 20.82, 95% CI 6.86, 63.21 for those born on 15th July 2015), however the direction of this effect reversed over the course of the trial, with those born in a clinic more likely to be tested (aOR 0.43, 95% CI 0.27, 0.70 for those born on 15th December 2015; aOR 0.06, 95% CI 0.04, 0.10 for those born on 15th July 2016; and aOR 0.14, 95% CI 0.04, 0.44 for those born on 15th December 2016). There was no evidence of a change in the difference

Figure 6.2 - Effect of delivery location on PCR testing over calendar time, based on data from MONARCH
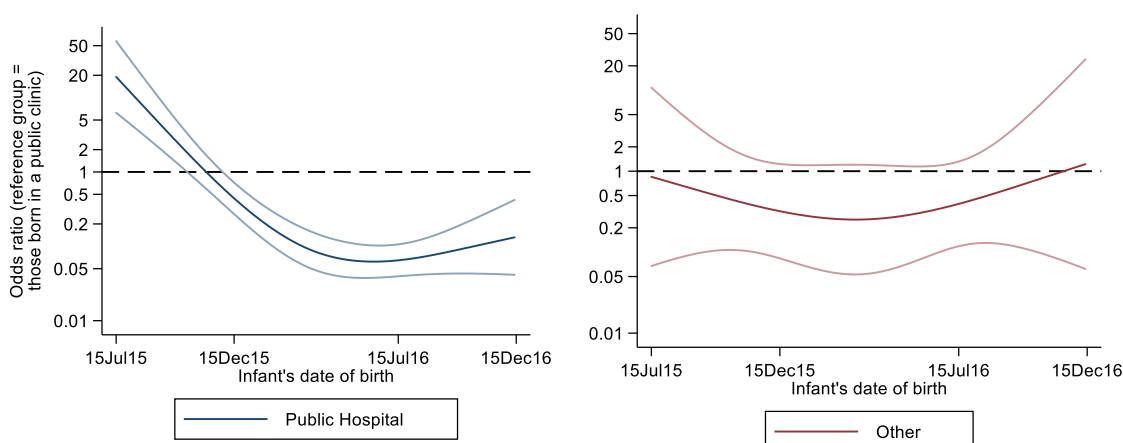
in testing coverage among those born in other locations compared to a clinic, although this group included only a small number of individuals.

6.5. Discussion

In the analysis using NHLS-AHRI data, infants born to women who had initiated ART prior to delivery were more likely to receive a PCR test than those born to women who were not. This association has also been observed in other studies in low- and middle-income settings [242, 243], despite it contradicting guidelines which prioritise testing in infants born to women recently initiating treatment and at higher risk of transmission. Women who are better engaged in care themselves may have increased understanding of their HIV status and the potential for transmission, and there may be more opportunities for healthcare workers to identify their infants as in need of HIV testing. ART use may have been misclassified for some mothers, given that it was based on linkage between ACDIS and ACCDB, which would have led to the underestimation of its effect size. A similar trend was observed in MONARCH, although the difference was not statistically significant, which may have been the result of the smaller sample size.

In the MONARCH dataset, there was some evidence of increasing testing coverage among women with a higher number of previous live-born children, even after adjustment for age. The increase did not appear to continue among the group with 3 or more previous children, although there was a wide confidence interval around this estimate, as a result of the smaller number of individuals. Possible mechanisms for this effect include a better understanding of infant testing processes among women who already have children. There was no evidence of this effect in the analysis based on NHLS-AHRI data, although, other than chance, there are no clear reasons why this would be.

As described in Chapter 5, using data from NHLS-AHRI surveillance, testing coverage initially increased over time, before decreasing again from 2013 onwards. It is possible this was in part a consequence of the change to national testing guidelines with the introduction of birth testing in 2015, but may also be explained by NHLS PCR test data being missing from one clinic for some years. Over the course of the MONARCH trial testing coverage decreased, with the effect appearing to be driven by a reduction in coverage among those born in a hospital rather than in a clinic or another location. Differences may also be caused by the difference in the definition of the outcome, with PCR test data from MONARCH only available to the 6 week postnatal visit.

No effect of delivery location was observed in the NHLS-AHRI surveillance analysis. This may have been caused by differences in the inclusion criteria between the two studies. Only those

living in the AHRI surveillance area were included in the NHLS-AHRI analysis, however all women delivering in the area were included in MONARCH, regardless of the location of their place of residence. Women may travel longer distances to deliver in a hospital compared to a clinic, meaning more of the infants born in the hospital may live outside the surveillance area and therefore have been tested elsewhere, which would not be captured here.

Multiple imputation was considered as an alternative method for dealing with the high proportion of missing data for some of the covariates in each dataset. Estimates from the complete case analysis were unbiased and depending on the missingness mechanism, estimates following multiple imputation may also have been, although the use of multiple imputation may have led to a more efficient analysis, with smaller standard errors for the estimated regression coefficients [231]. However, the majority of mother-infant pairs were included (>90% in each analysis), so efficiency gains would not be large. Multiple imputation would also have enabled exploration of the effect of the covariates that had to be excluded from the complete case analysis, such as CD4 and viral load, although given the very high proportion of missing data for many of these covariates, any imputed values generated may not be reliable. In fitting the GEE model with an exchangeable covariance structure, it was assumed that the correlation between all of the infants born to each mother was equal; this may not have been the case in situations where women had a multiple pregnancy (i.e. twins or triplets) as well as a singleton pregnancy.

Few of the factors investigated here were associated with PCR testing, even in univariable analyses. This may suggest that much of the likelihood of receiving a PCR test depends on processes within the clinic to identify HIV-exposed infants, rather than any individual characteristics about the infant or their mother. Alternatively, the misclassification of PCR testing status, which seems likely given the differences in estimated coverage between the two datasets reported in Chapter 5, would have resulted in the attenuation of any true effects. As an example, there was no strong evidence of an effect of maternal age here, although this has been reported in other studies [244]. As well as variables which were excluded because of a high proportion of missing data, there are several of potential variables of interest which could not be explored as no data were available on them; previous studies have shown association with maternal status disclosure, depression score, social support and knowledge of EID [242, 244].

6.6. Key findings

The key findings from this chapter are:

- Testing coverage increased over time to 2013, then subsequently decreased again
- Infants whose mothers who had ever initiated ART were more likely to receive a PCR test compared to those whose mothers were not
- In the MONARCH dataset only, after the first 6 months of the trial, there was evidence that infants born in a hospital were less likely to receive a PCR test than those born elsewhere

Chapter 7. Frequency, timing and results of HIV PCR testing

7.1. Introduction

In this chapter, I investigate the frequency, timing and results of PCR testing of HIV-exposed infants, among those who ever received a PCR test, using the NHLS PCR data. Given that the recommended time points for PCR testing coincide with those for vaccinations, coverage of PCR testing was compared to coverage of vaccinations at each timepoint, using vaccination data from AHRI's demographic surveillance database, ACDIS, both at an individual- and population-level.

7.2. Objectives

The objectives for this chapter are to:

- Estimate the proportion of infants tested in accordance with the South African guidelines at the time of their birth
- Estimate the proportion of infants testing HIV PCR positive
- Use vaccination data collected in demographic surveillance to compare overall engagement in care with the coverage of HIV PCR testing

7.3. Methods

All infants born between 1st June 2010 and 31st December 2016 who received an HIV PCR test conducted at one of the 17 healthcare facilities within the Hlabisa health sub-district were included in this analysis, with PCR test data available from the NHLS database and unique infants within the database identified through the linkage process described in Chapter 4.

The number of PCR tests received by each infant and the proportion of infants receiving tests in accordance with the guidelines at the time of their birth were described, both overall and by testing guideline period (<1st April 2015 vs. ≥1st April 2015). Prior to April 2015, guidelines had recommended a test at 6 weeks of age, but they were subsequently updated to recommend a test at birth, and either at 10 weeks of age or at 18 weeks of age if the infant received extended 12 week nevirapine prophylaxis. An additional test at 6 weeks after the cessation of breastfeeding was recommended throughout both time periods. In this analysis, a window of +1 week was used to identify tests at birth, and a window of ±2 weeks was used to identify tests at other time points. As no data on nevirapine prophylaxis use were available to determine the optimal timing of the follow-up test for infants born after the introduction of birth testing, any

test between 10 and 18 weeks of age (plus a window of ±2 weeks) was considered compliant to the guidelines. Among those born after 1st April 2015, logistic regression was used to investigate whether both birth testing and follow-up testing after a negative test at birth improved over time after the change in guidelines, with date of birth included as a continuous variable.

The results of the PCR tests were summarised by age at test and by testing guideline time period, both for all tests and for the first test per infant only. The age of the infant at each PCR test conducted was classified as <4 weeks, ≥4-≤8 weeks and >8 weeks for those born prior to the introduction of birth testing, and as ≤1 week, >1-<8 weeks, ≥8-≤20 weeks, and >20 weeks for those born after, in line with the age at recommended tests (and allowing for windows around these).

The proportion of infants who ever received a positive PCR test was summarised by year of birth. Age at HIV diagnosis, defined as the age at first positive test, was summarised, along with the proportion of those ever testing positive who received a confirmatory HIV PCR test (among those born after 1st April 2015 only, as prior to this the national guidelines recommended the use of a viral load test to confirm HIV status). The proportion of infants who were retested following an indeterminate test result (among all infants regardless of date of birth) was also reported. For infants with more than one indeterminate test result, retesting after the last indeterminate test was considered. Due to small numbers, changes in confirmatory testing and retesting following an indeterminate result over time were not explored.

To assess the potential impact of failing to link repeat PCR tests on the same infants on the estimate of confirmatory testing a sensitivity analysis was conducted. Unique infants born after the introduction of birth testing who ever received a positive PCR test and who shared their date of birth and sex (considered to be the most robust identifying variables) with another such infant were identified, excluding those with PCR tests conducted on the same day (who were assumed to be separate individuals). The estimate of the proportion of infants with a positive PCR test who received a confirmatory test was recalculated, assuming these infants were in fact the same child.

In order to explore variation in PCR testing, the following measures were summarised by clinic:

i)   the proportion of infants with a negative test result at birth who received a follow-up test (among those born after the introduction of routine birth testing on 1st April 2015), where results were summarised according to the clinic at which the birth test was conducted

ii)   the proportion of infants with a positive PCR test who received a confirmatory tests (among those born after 1st April 2015 only), summarised by the clinic at which the initial positive test was conducted

iii)   the proportion retested following an indeterminate test result, where results were summarised according to the clinic at which the indeterminate test was conducted.

As the clinic at which each infant was born was not available, the variation in coverage of birth testing by clinic could not be assessed. Proportions were compared by clinic using a chi-squared test, or Fisher's exact test where the number of infants from any clinic was fewer than 5. As described in Chapter 4, PCR test data from one clinic outside the sub-district were also available from NHLS, however, infants who were only tested at this clinic were excluded from analysis, making the assumption that they were born outside of the sub-district. Infants who were tested at birth at this clinic were therefore excluded from the analysis of follow-up testing after a negative test at birth, as by definition they had to have received at least one subsequent test.

Vaccination data, collected as part of AHRI's population surveillance, were used to compare PCR testing estimates with overall engagement in care during the first year of life. The South African vaccination schedule is shown in Table 1.11. A vaccination form is completed during annual surveillance for all children aged less than 5 years old who are resident within the AHRI demographic surveillance area. The form collects data on whether each vaccination was given (with tick boxes for each of the possible responses 'yes', 'no', and 'don't know'), and the date the vaccination was given if known. The form asks whether the Road-to-health booklet was seen at the time of data collection, with responses based on verbal reports from caregivers if not. The data collected on the form have changed over time; the questions about polio vaccinations from 10 weeks of age onwards were dropped after 2015, as this vaccination was given in combination with DTP+HIB [245].

Among children born since 1st June 2010 and resident in the AHRI demographic surveillance area before the age of 5 years, the proportion who received each vaccination was calculated, with children excluded from the analysis of any vaccinations due after the last time they had a vaccination form completed. If all responses for a child were 'don't know' or were missing then it was assumed that the form had not actually been completed, and the child was excluded from all analyses. Among the remaining children, the proportion who received each vaccination was calculated using three different methods. Method 1 assumed that a given vaccination was not received if the response on the vaccination form was unknown or missing. Two further definitions were applied as a sensitivity analyses; method 2 excluded those with unknown/missing response for each vaccination and method 3 excluded those with

unknown/missing response, unless their Road-to-health-booklet was reported to have been seen at the time of form completion, in which case it was assumed the vaccination had not been received. Estimates of coverage calculated using data from ACDIS were compared to those made for South Africa by both WHO-UNICEF (United Nations International Children's Emergency Fund) and the Department of Health, where available [246].

In addition to looking at the coverage of each vaccination separately, the overall vaccination coverage was estimated, defined as the proportion of infants who received all recommended vaccinations to 1 year of age. Vaccinations with 'missing' responses were excluded from the denominator for each infant, with 'missing' defined using each of the three methods described above. Estimates of overall coverage were compared to those published in the District Health Barometer reports for both South Africa and KwaZulu-Natal [247].

Finally, for the subset of infants known to be HIV-exposed (whose mother was identified as having HIV either through the serosurvey and/or linkage to ACCDB, see method 1 in Chapter 5) and who were resident within the demographic surveillance area at birth, PCR testing coverage and vaccination coverage were compared (using linkage between the NHLS data and ACDIS) at the relevant time points according to the guidelines at the time of their birth using chi-squared tests, with method 1 used to deal with missing data when determining which vaccinations were received.

7.4. Results

Within the NHLS PCR test data, 15,234 unique infants were identified, with a total of 17,622 PCR tests conducted. 10,490 (69%) infants were born before the introduction of birth testing on 1st April 2015, and 4,744 (31%) were born on or after 1st April 2015.

Number and timing of tests among those born before 1st April 2015 (before the introduction of birth testing)

Among the 10,490 infants born prior to the introduction of birth testing and ever PCR tested, 9,167 (87%) received 1 test, 1,221 (12%) received 2 tests, 92 (1%) received 3 tests, and 10 (<0.5%) received 4 tests. The age of the infant at each PCR test conducted is shown in Figure 7.1, with the majority occurring at either 6 or 7 weeks of age.

Figure 7.1 - Age in weeks at each PCR test conducted among infants born before 1ˢᵗ April 2015
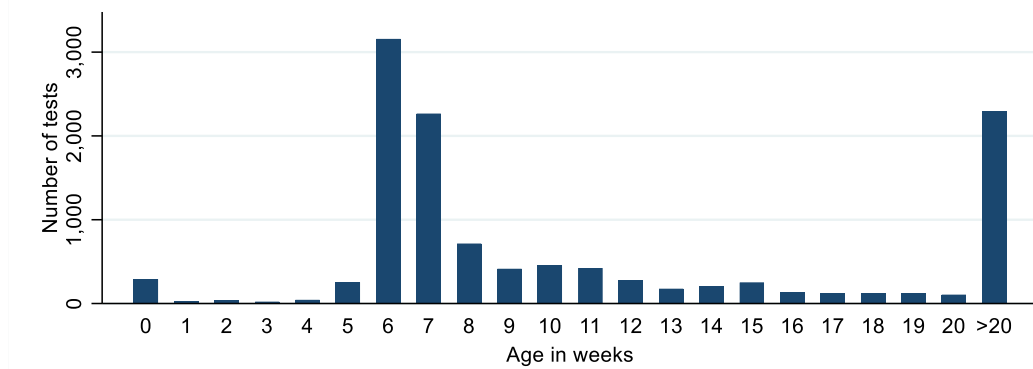
The median (IQR) age at first test was 7.0 (6.1, 11.1) weeks, and 925 (9%) and 219 (2%) infants received their first test when aged over 6 months and 1 year respectively. 6,208 (59%) infants received a test at 6 weeks (allowing a window of ±2 weeks) of age as recommended.

Number and timing of tests among those born on or after 1ˢᵗ April 2015 (after the introduction of birth testing)

Among 4,744 infants born after the change to the testing guidelines and ever tested, 3,917 (83%) received 1 test, 717 (15%) received 2 tests, 95 (2%) received 3 tests and 15 (<0.5%) received 4 or more, resulting in a total of 5,697 tests. There was a statistically significant increase in the proportion of infants with more than one test compared to prior to the introduction of birth testing (17% vs. 13%, p<0.001). Of 827 infants with more than one test, 183 (22%) were tested at more than one clinic.

The age of the infant at each PCR test conducted is shown in Figure 7.2. The median (IQR) age at each PCR test was 6.0 (0.0, 11.7) weeks; 2,610 (46%) tests were conducted at birth, 698 (12%) between 1 and 7 weeks, 1,745 (31%) between 8 and 20 weeks of age (allowing a window of ±2 weeks around the recommended time for follow-up tests of 10-18 weeks), and 644 (11%) after 20 weeks of age.

Figure 7.2 - Age in weeks at each PCR test conducted among infants born after 1st April 2015
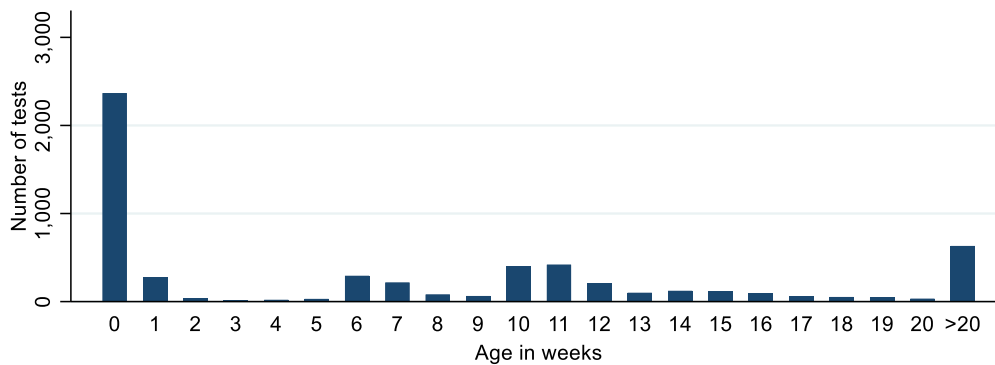
Of 4,744 infants ever tested, 2,588 (55%) were first tested at birth, of whom 1,578 (61%) were tested on the day they were born, 772 (30%) were tested aged 1 to 3 days old, and the remaining 238 (9%) were tested between 4 and 7 days old. Of those testing negative at birth (n=2,575), 576 (22%) received a follow-up test at a median (IQR) age of 10.9 (9.4, 14.6) weeks, and 393 (15%) received a second test at 10-18 weeks of age as recommended. Among infants not tested at birth (n=2,156), the first test was at a median age of 10.9 (7.4, 15.3) weeks, with 246 (5% of all those born after the introduction of birth testing) not tested until after 6 months of age. Those not tested at birth were significantly more likely to receive a test at 10-18 weeks of age than those with a negative test at birth (1,295/2,156 (60%) vs. 393/2,575 (15%), p<0.001).

The proportion of infants receiving a test at birth, and receiving a follow-up test following a negative result at birth is shown by month of birth in Figure 7.3. Coverage of birth testing was much lower among those born in April 2015, at 15%, compared to 47%-65% by month of birth for those born later. From May 2015 onwards, there was some evidence of an increase in odds of being tested at birth for each month of birth later an infant was born (odds ratio (OR) 1.01; 95% CI 1.00, 1.02; p=0.072). Conversely, coverage of follow-up testing following a negative result at birth was much higher for those born in April 2015 than in subsequent months (56%, compared to 16-27% by month for those born later), but from May 2015 onwards there was no evidence of this changing over time (OR 1.01; 95% CI 0.99, 1.02; p=0.408).

Among 45 infants who tested negative at birth, never received a subsequent PCR test and were linked to ACDIS, none were known to have died.

Figure 7.3 - Percentage of infants tested at birth and percentage of infants who received a follow-up test following a negative test at birth, by month of birth
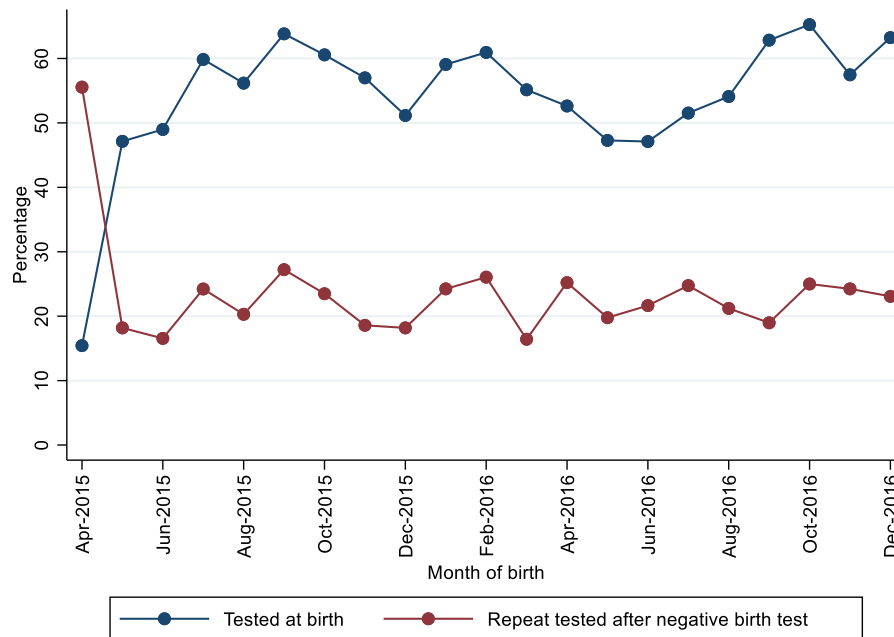
PCR test results

Results of PCR tests conducted are shown in Table 7.1 by age at test and guideline time period, both for all tests together and restricted to the first test per infant only. In both time periods, the proportion of all tests which were positive was lower among tests conducted at ages recommended under the guidelines (≥4-≤8 weeks before the introduction of birth testing, and ≤1 week and ≥8-≤20 weeks after the introduction) compared to those conducted outside of this (p<0.001, p<0.001).

Overall, the proportion of infants who ever received a positive test result was 3.0% (458/15,234), and fell from 6.6% (69/1,044) among those born in 2010 to 1.0% (24/2,523) among those born in 2016 (Figure 7.4). A similar trend was observed in the proportion testing positive by one year of age. The proportions testing positive were 3.7% (390/10,490) and 1.4% (68/4,744) among those born before and after the introduction of birth testing respectively. The median age at first positive test result was 10.6 (6.7, 27.4) weeks among those born prior to the introduction of birth testing and 13.1 (2.3, 28.6) weeks among those born after, representing a non-statistically significant increase (p=0.400).

Table 7.1 - PCR test results by age at test and by testing guideline time period, for all tests and for the first test per infant only

Guideline time period	Age at test		Positive	Negative	Indeterminate
Prior to introduction of birth testing (N=10,490)	First test per infant	<4 weeks (n=389)	23 (5.9%)	359 (92.3%)	7 (1.8%)
		≥4-≤8 weeks (n=6,197)	**136 (2.2%)**	**5,964 (96.3%)**	**97 (1.6%)**
		>8 weeks (n=3,904)	205 (5.3%)	3,625 (92.9%)	74 (1.9%)
	All tests (N=11,925)	<4 weeks (n=433)	25 (5.8%)	399 (92.2%)	9 (2.1%)
		≥4-≤8 weeks (n=7,166)	**163 (2.3%)**	**6,887 (96.1%)**	**116 (1.6%)**
		>8 weeks (n=4,326)	228 (5.3%)	4,009 (92.7%)	89 (2.1%)
After introduction of birth testing (N=4,744)	First test per infant	**≤1 week (n=2,588)**	**13 (0.5%)**	**2,554 (98.7%)**	**21 (0.8%)**
		>1-<8 weeks (n=575)	12 (2.1%)	557 (96.9%)	6 (1.0%)
		≥8-≤20 weeks (n=1,240)	14 (1.1%)	1,222 (98.6%)	4 (0.3%)
		>20 weeks (n=341)	17 (5.0%)	320 (93.8%)	4 (1.2%)
	All tests (N=5,697)	**≤1 week (n=2,610)**	**13 (0.5%)**	**2,574 (98.6%)**	**23 (0.9%)**
		>1-<8 weeks (n=698)	18 (2.6%)	672 (96.3%)	8 (1.2%)
		≥8-≤20 weeks (n=1,745)	**29 (1.7%)**	**1,708 (97.9%)**	**8 (0.5%)**
		>20 weeks (n=644)	41 (6.4%)	594 (92.2%)	9 (1.4%)

Recommended tests: at 6 (±2) weeks prior to the introduction of birth testing, and at birth (+1 week) and 10-18 (±2) weeks after the introduction of birth testing. Rows in bold represents recommended timepoints for tests

Figure 7.4 - Percentage of infants ever receiving a positive PCR test result and receiving a positive PCR test by one year of age, by year of birth

164
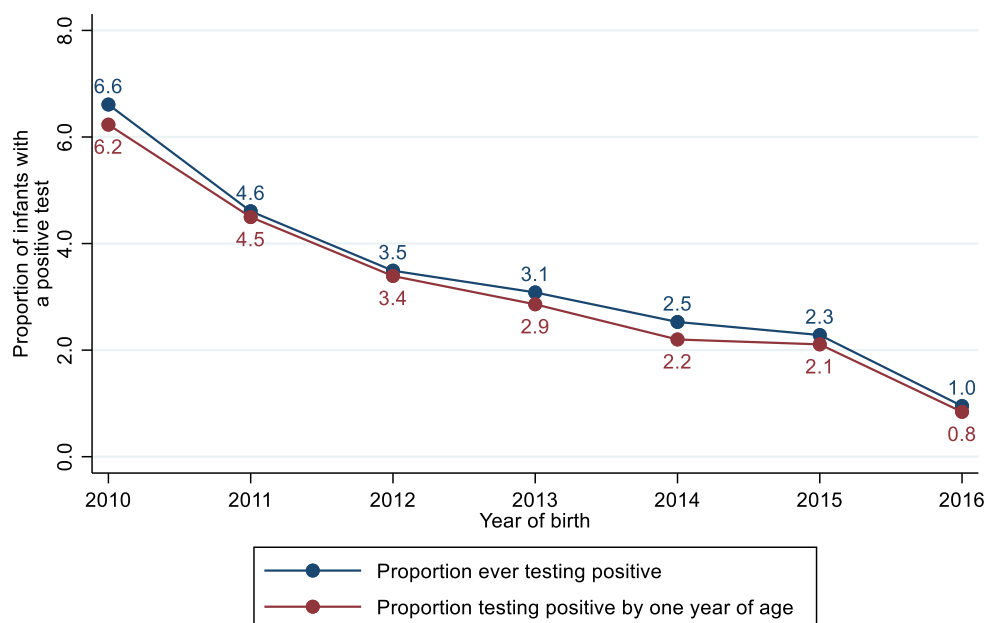
Among the 68 infants born after the introduction of birth testing and ever testing positive, 13 (19%) first tested positive at birth, corresponding to a yield of 0.5% (13/2,588) of all birth tests, 7 (10%) infants initially tested negative at birth but were subsequently diagnosed at a median (IQR) age of 15.8 (7.7, 21.7) weeks, and the remaining 48 (71%) infants were not tested at birth but were diagnosed at a median age of 16.5 (7.3, 33.9) weeks.

Of the 68 infants ever testing positive during this time period, 30 (44%) received a confirmatory test at a median of 13 (7, 19) days later. All these infants were confirmed to have HIV, apart from one (aged 1.5 years at last PCR test) for whom all subsequent test results were indeterminate. Of the 68 infants, 2 had the same date of birth and sex; assuming these 2 infants were in fact the same individual, the proportion of diagnosed infants with a confirmatory test increased to 46% (31/67).

Among all infants regardless of date of birth, 242 ever received an indeterminate test result, of whom 18 infants received more than one, corresponding to 1.5% (262/17,622) of all PCR tests conducted. The proportion of infants who ever received an indeterminate test result is shown in Figure 7.5 by year of birth, and varied from 4.4% (91/2,084) in 2011 to 0.6% (16/2,523) in 2016. Following the indeterminate test result, 137 infants (57%) were retested at a median (IQR) 32 (26, 64) days later, of whom 10 (7%) tested positive and 126 (93%) tested negative. Of the 99 (43%) with no subsequent test, the median age at indeterminate test was 8.4 (5.7, 33.1) weeks, and 13 (13%) were older than 1 year of age. There was some evidence that repeat testing was higher prior to the introduction of birth testing compared to after the change in April 2015 (117/198 (59%) vs. 20/44 (45%), p=0.099).

PCR testing by clinic

The results of testing by clinic are shown in Table 7.2. The proportion of infants who ever received a follow-up test after testing negative at birth varied between clinics from 7% to 58% (p<0.001), the proportion with a confirmatory test following a positive result varied from (among clinics with ≥5 positive tests) between 17% and 80% (p=0.140), and retesting following an indeterminate result varied (among clinics with ≥5 indeterminate tests) from 44% to 88% (p=0.013).

Figure 7.5 - Proportion of infants ever receiving an indeterminate PCR test result, by year of birth
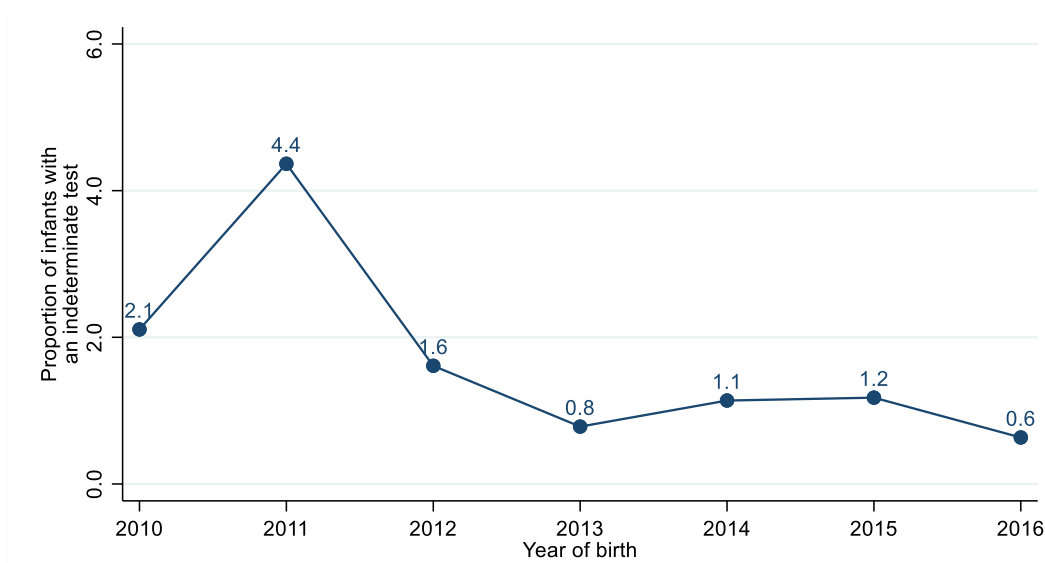
Table 7.2 - Repeat PCR testing by clinic

Clinic	Follow-up testing after negative test at birth (infants born after 1st April 2015)	Confirmatory testing after positive test (infants born after 1st April 2015)	Retesting after indeterminate test (all infants)
1	3/14 (21%)	-	0/1 (0%)
2	3/18 (17%)	1/2 (50%)	-
3	14/67 (21%)	0/1 (0%)	4/8 (50%)
4	80/786 (10%)	8/20 (40%)	25/55 (45%)
5	25/141 (18%)	1/6 (17%)	12/18 (67%)
6	90/342 (26%)	2/3 (67%)	25/55 (45%)
7	55/118 (47%)	0/1 (0%)	10/12 (83%)
8	9/41 (22%)	1/3 (33%)	-
9	39/104 (38%)	4/4 (100%)	6/10 (60%)
10	36/62 (58%)	2/3 (67%)	5/11 (45%)
11	5/59 (8%)	0/2 (0%)	1/4 (25%)
12	30/112 (27%)	5/7 (71%)	5/10 (50%)
13	42/183 (23%)	1/4 (25%)	6/10 (60%)
14	33/139 (24%)	1/5 (20%)	4/9 (44%)
15	7/33 (21%)	0/1 (0%)	3/4 (75%)
16	78/214 (36%)	0/1 (0%)	23/26 (88%)
17	8/123 (7%)	4/5 (80%)	7/10 (70%)
Overall	557/2,556 (22%)	30/68 (44%)	136/242 (56%)
p	<0.001	0.140	0.013
p (including clinics with ≥5 tests only)	<0.001	0.096	0.022

p-values compare the proportions by clinic for each measure using chi-squared tests (where all clinics have ≥5 tests) or Fisher's exact test. 19 infants with a negative test at birth at a clinic outside the sub-district are excluded from the analysis of follow-up testing following a negative test at birth.

The difference from the overall proportion for each measure for each clinic is shown in Figure 7.6. Some clinics were better than average (e.g. clinic 9) or worse than average (e.g. clinic 11)

across all three measures, although for the majority performance varied by measure. Repeat testing did not appear to be strongly related to clinic size.

Figure 7.6 - Difference in PCR testing from overall average by clinic
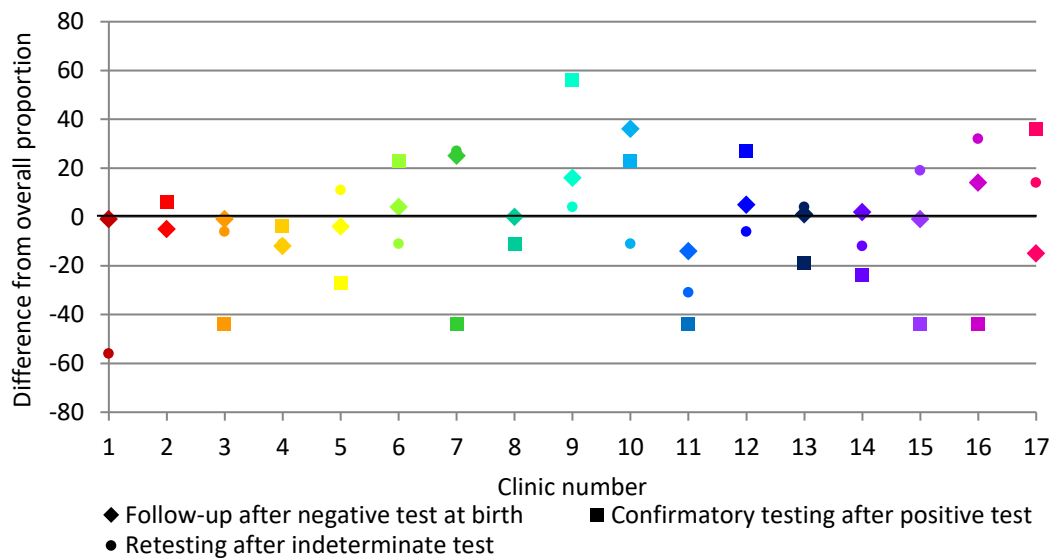
Comparison between PCR test and vaccination data

Among 11,541 children in ACDIS who were born since 1st June 2010 and were resident in the demographic surveillance area before the age of 5 years, 8,726 (77%) had a vaccination form completed. Of those who ever had a vaccination form completed, all had a form completed after 14 weeks of age, 8,516 (98%) after 9 months of age, 7,682 (88%) after 18 months of age, and 1,528 (18%) after 5 years of age. The Road-to-health booklet was not seen at the time of data collection for most children (7,700, 88%), although 99% (n=8,617) of children were reported to have one.

Table 7.3 - Vaccination coverage by vaccine

Age of child		Number with a form completed past this age	Proportion receiving each vaccination			If received, is date available?	Estimate of national coverage [246]	
			Method 1	Method 2	Method 3		WHO - UNICEF	DoH
Birth	BCG	8,726	8,553/8,726 (98%)	8,553/8,586 (100%)	8,553/8,590 (100%)	3,224 (38%)	67%	83%
	Polio	8,726	8,559/8,726 (98%)	8,559/8,591 (100%)	8,559/8,595 (100%)	3,220 (38%)		
6 weeks	Polio	8,726	8,524/8,726 (98%)	8,524/8,576 (99%)	8,524/8,579 (99%)	3,123 (37%)		
	DTP+HIB	8,726	8,534/8,726 (98%)	8,534/8,576 (100%)	8,534/8,579 (99%)	3,137 (37%)	83%	94%
	Hepatitis B	8,726	8,531/8,726 (98%)	8,531/8,573 (100%)	8,531/8,576 (99%)	3,133 (37%)		
	Rotavirus	8,726	8,333/8,726 (95%)	8,333/8,511 (98%)	8,333/8,514 (98%)	3,029 (36%)		
	PCV	8,726	8,325/8,726 (95%)	8,325/8,510 (98%)	8,325/8,514 (98%)	3,008 (36%)		94%
10 weeks	Polio*	6,248	5,871/6,248 (94%)	5,871/6,024 (97%)	5,871/6,045 (97%)	1,861 (32%)		
	DTP+HIB	8,726	8,464/8,726 (97%)	8,464/8,557 (99%)	8,464/8,562 (99%)	3,054 (36%)		
	Hepatitis B	8,726	8,453/8,726 (97%)	8,453/8,556 (99%)	8,453/8,562 (99%)	3,045 (36%)		
14 weeks	Polio*	6,225	5,868/6,225 (94%)	5,868/5,988 (98%)	5,868/6,008 (98%)	1,916 (33%)	75%	93%
	DTP+HIB	8,726	8,407/8,726 (96%)	8,407/8,520 (99%)	8,407/8,527 (99%)	3,025 (36%)	75%	93%
	Hepatitis B	8,726	8,386/8,726 (96%)	8,386/8,502 (99%)	8,386/8,509 (99%)	3,012 (36%)	75%	93%
	Rotavirus	8,726	8,125/8,726 (93%)	8,125/8,423 (96%)	8,125/8,437 (96%)	2,850 (35%)	77%	95%
	PCV	8,726	8,100/8,726 (93%)	8,100/8,401 (96%)	8,100/8,419 (96%)	2,840 (35%)		94%
9 months	Measles	8,516	7,470/8,516 (88%)	7,470/8,008 (93%)	7,470/8,056 (93%)	2,422 (32%)	76%	97%
	PCV	8,516	7,272/8,516 (85%)	7,272/7,934 (92%)	7,272/7,989 (91%)	2,310 (32%)	77%	95%
18 months	Polio*	5,118	3,404/5,118 (67%)	3,404/4,554 (75%)	3,404/4,612 (74%)	660 (19%)		
	DTP	7,682	5,496/7,682 (72%)	5,496/6,650 (83%)	5,496/6,731 (82%)	1,388 (25%)		79%
	Measles	7,682	5,437/7,682 (71%)	5,437/6,579 (83%)	5,437/6,661 (82%)	1,381 (25%)	66%	84%
5 years	Polio*	737	20/737 (3%)	20/470 (4%)	20/490 (4%)	1 (5%)		
	DT	1,528	30/1,528 (2%)	30/681 (4%)	30/745 (4%)	5 (17%)		

DoH: Department of Health; DT: Diphtheria and tetanus; DTP+HIB: Diphtheria, tetanus, acellular pertussis, and haemophilus influenza type B; UNICEF: United Nations International Children's Emergency Fund; WHO: World Health Organization

*From December 2015 questions about the polio vaccine from 10 weeks of age onwards was dropped from the vaccination form as this was given in combination with DTP+HIB

The proportion of infants receiving each vaccination, estimated using each of the three methods for dealing with missing data, is shown in Table 7.3. Using method 1, coverage of each of the recommended vaccinations to 14 weeks of age was at least 93%, falling to 85-88% for vaccinations at 9 months, 67-72% at 18 months, and 2-3% for those recommended at 5 years of age. Estimates of coverage made using both method 2 and method 3 were relatively similar to those from method 1. National estimates of coverage made by the Department of Health were generally comparable to those estimated here, although those made by WHO-UNICEF were approximately 20% lower at each timepoint.

The proportion meeting the definition of overall coverage is shown in Table 7.4. Overall coverage was estimated to be 78% using method 1, ranging by year of birth from 69% (among those born in 2015) to 82% (among those born in 2011) with no clear trend over time. Again, the estimates of coverage calculated using methods 2 and 3 were relatively similar.

Table 7.4 - Overall vaccination coverage, by year of birth

Year of birth	Method 1	Method 2	Method 3	District Health Barometer report estimates [247]	
				SA	KZN
2010 (n=1,172)	836 (71%)	936 (80%)	933 (80%)	81%	78%
2011 (n=1,992)	1,587 (80%)	1,708 (86%)	1,702 (85%)	84%	88%
2012 (n=1,817)	1,497 (82%)	1,639 (90%)	1,624 (89%)	84%	86%
2013 (n=1,517)	1,243 (82%)	1,372 (90%)	1,358 (90%)	84%	86%
2014 (n=1,239)	975 (79%)	1,118 (90%)	1,108 (89%)	90%	90%
2015 (n=832)	575 (69%)	694 (83%)	674 (81%)	89%	85%
2016 (n=157)	117 (75%)	127 (81%)	125 (80%)	82%	86%
Total (n=8,726)	6,830 (78%)	7,594 (87%)	7,524 (86%)	-	-

SA: South Africa; KZN: KwaZulu-Natal
Overall vaccination coverage defined as the proportion of infants receiving all recommended vaccinations to 1 year of age, excluding those with no data available (defined according to each method)

In total, 1,734 HIV-exposed and 6,992 HIV-unexposed infants had a completed vaccination form. A comparison of the coverage of PCR testing and of vaccinations (using vaccination coverage method 1) among these infants is shown in Table 7.5.

Among infants born before 1st April 2015, 79% (565/713) of HIV-exposed infants who ever received a PCR test and 76% (623/815) of those who never received a PCR test received all their vaccinations to 1 year of age, compared to 80% (5,155/6,483) of HIV-unexposed infants born in same time period. Among the HIV-exposed infants, only 2% (19/1,029) of those with no PCR test at 6 weeks of age also did not receive their 6 week vaccinations, compared to 3% (15/499) of those who did receive a PCR test (p=0.150).

Among infants born after 1st April 2015, 67% (47/70) of HIV-exposed infants who ever received a PCR test and 70% (95/136) of those who never received a PCR test received all their

vaccinations to 1 year of age, compared to 68% (345/509) of HIV-unexposed infants born in same time period. Only 1% (2/184) of HIV-exposed infants with no PCR test at birth also did not receive any vaccinations at birth (compared to 0% (0/22) of those who did receive a birth test, p=0.623), and 4% (6/170) of those with no PCR test at 10-18 weeks did not receive their 10 week vaccinations (compared to 0% (0/36) of those who did receive a PCR test, p=0.253).

Across both time periods, there was no difference in the proportion of infants who received all vaccinations to 1 year of age among those diagnosed with HIV (21/28, 75%) compared to those not (596/760, 78%) (p=0.666).

Table 7.5 - A comparison of the coverage of PCR testing and vaccinations

Time period	Vaccination coverage	HIV exposed (N=1,734)			HIV unexposed (N=6,992)
		Received PCR test	Did not receive PCR test	p	
Prior to introduction of birth testing	At 6 weeks — Received vaccination	484 (97%)	1,010 (98%)		6,353 (98%)
	Did not receive vaccination	15 (3%)	19 (2%)	0.150	130 (2%)
	Total	499	1,029		6,483
	All to 1 year of age — Received all vaccinations	565 (79%)	623 (76%)		5,155 (80%)
	Did not receive vaccinations	148 (21%)	192 (24%)	0.189	1,328 (20%)
	Total	713	815		6,483
After introduction of birth testing	At birth — Received vaccination	22 (100%)	182 (99%)		504 (99%)
	Did not receive vaccination	0	2 (1%)	0.623	5 (1%)
	Total	22	184		509
	At 10-18 weeks — Received vaccination	36 (100%)	164 (96%)		499 (98%)
	Did not receive vaccinations	0	6 (4%)	0.253	10 (2%)
	Total	36	170		509
	All to 1 year of age — Received all vaccinations	47 (67%)	95 (70%)		345 (68%)
	Did not receive vaccinations	23 (33%)	41 (30%)	0.691	164 (32%)
	Total	70	136		509

7.5. Discussion

Among infants born after the introduction of birth testing and ever tested, only about half actually received a test at birth. This is much lower than another estimate of birth testing coverage from South Africa, which reported an increase in national coverage from 39% in June 2015 (when birth testing was introduced in provinces other than KwaZulu-Natal) to 93% in May 2016, with coverage higher than average in KwaZulu-Natal [122]. The authors of this study were unable to identify repeat PCR tests on each individual in their data, and so the numerator used was the number of tests conducted in the first 7 days of life. This would have led to an overestimation of coverage, although the number of infants receiving multiple tests in the first week of life may be expected to be small (in my analysis it was 2%). As their estimate included all exposed infants in the denominator, rather than restricting to just those ever tested as has been done here, it would be expected to be lower than the estimate calculated for this analysis. There was evidence of improvement in the coverage of birth testing over time after the change

in guidelines, consistent with results here; initial low coverage may have been a result of a lack of awareness or confusion over new guidelines.

Following a negative test at birth, less than a quarter of infants in this analysis ever received another PCR test. As a result of the scale up of PMTCT interventions most infants with perinatal infection now acquire HIV in utero rather than intrapartum [49], however failure to retest infants could still result in missed intrapartum infections, with the sensitivity of HIV PCR tests only estimated to be 55% at birth [46], highlighting the importance of follow-up testing. This is comparable to the national study mentioned above which estimated repeat testing to increase from 13% to 48% between June 2015 and May 2016, although here repeat testing was estimated based only on the number of tests conducted at 10 weeks of age rather than actually linking tests at an individual level [122]. Estimates from studies conducted in urban areas were much higher, with coverage closer to 80% [160, 164], as were those looking at testing following a targeted birth test, which may be expected due to the higher risk of perinatal transmission in this population [161, 165]. No data to determine the risk status of infants were available for my analysis to enable the calculation of comparable estimates. Of note, data from the MONARCH trial, implemented in clinics in the sub-district, also showed sub-optimal coverage of repeat testing among women initially presenting without HIV to antenatal care, with only 67% receiving a further test at any point in pregnancy (as recommended in PMTCT guidelines) [138].

Throughout the whole time period, guidelines recommended an additional HIV test at 6 weeks after the cessation of breastfeeding, which should be a PCR test if the infant is still aged less than 18 months, or an antibody test otherwise. Although no data on breastfeeding were available to assess adherence to this at an individual level, there were few tests conducted in older infants, suggesting it was not commonly done. Since it was estimated that in 2016 47% of infants in South Africa were still breastfed between the age of 12 and 17 months [110], it is likely that for many of the infants included in this analysis it would have been age-appropriate to confirm their HIV status using an antibody test [121], rather than a PCR test. The results of antibody tests are not routinely collected anywhere to assess this. Of the infants who were tested at older ages, a higher proportion tested positive, suggesting this testing may have been targeted at symptomatic or high risk infants.

Importantly, less than half (44%) of infants who tested positive in this analysis were observed to receive a subsequent PCR test confirming their HIV status. This is consistent with results from a study across KwaZulu-Natal, where only 31% of infants born in 2015 and testing positive received a confirmatory test [157]. Another study, which also used probabilistic data linkage to identify repeat tests within the NHLS dataset, reported that 46% of infants testing positive received a confirmatory test [180]. Results from a South African modelling study concluded that without confirmatory testing more than 10% of infants who have one positive result and initiate ART

may be false-positive diagnoses [248], resulting in unnecessary cost, emotional impact, and potential ART toxicity. In addition, it is difficult to identify who these false-positives are after the initiation of ART, as infants who do have HIV would be expected to achieve an undetectable viral load, and may also be antibody negative if treatment was initiated early [249]. Although most infants in this analysis who received a confirmatory test again tested positive, all subsequent tests for one infant were indeterminate, complicating decisions about future care. The coverage of retesting following an initial indeterminate result was slightly better than following a positive test result, though still not optimal, at approximately 60%.

Very high vaccination coverage was observed in data collected during annual demographic surveillance, especially to 14 weeks of age, by which time most PCR tests should have been conducted. Linkage of infants between the surveillance and NHLS laboratory data showed that almost all of those who did not receive PCR tests at each of the relevant time points had actually attended clinic to receive routine vaccines at that time. Clinics within the sub-district operate parallel systems for HIV PCR testing and other routine infant check-ups (including vaccinations), with different nurses in different parts of the clinics responsible for each. These results suggest that although infants and their caregivers are engaged in care and are attending clinic, those in need of PCR tests are not successfully being identified once there. Reasons for this may include poor recording of maternal HIV status in the infant's Road-to-health booklet, or healthcare workers' reluctance to disclose maternal HIV status in the presence of other family members [250]. There was evidence of lower engagement with healthcare services at older ages with fewer infants reported to have received their vaccinations at 9 and 18 months of age, which may suggest lower coverage of PCR testing following the cessation of breastfeeding.

It is possible that vaccination coverage has been overestimated here as the majority of reports were based on recall by caregivers, with the Road-to-health booklet not seen for almost 90% of infants. It is possible that caregivers would report that their infant was vaccinated even if they were not (social desirability bias), although maternal recall has been shown to be accurate in a previous study in the sub-district (with no difference by maternal HIV status) [251]. Results from the sensitivity analyses using alternative definitions of coverage gave higher estimates when restricting the analysis to those with available data on each vaccination, as expected. In addition, infants with missing data on all vaccinations were excluded from analysis but may actually just have not been vaccinated, leading to the overestimation of vaccine coverage. Only three-quarters of eligible infants living within the surveillance had a vaccination form completed; coverage among these infants may not have been the same as among those who did have an available form. Of note, estimates of vaccination coverage here were higher than national estimates [246].

Although there were small numbers of patients at some clinics, there was some evidence of variation in the receipt of follow-up testing, confirmatory testing, and retesting following an indeterminate result, between clinics, although it was not always the same clinics which did better or worse across all measures. This variation may be further evidence that low testing coverage is at least partly the result of poor clinic procedures, as well as poor engagement in care. Differences in performance may provide an opportunity to identify more successful follow-up procedures.

The proportion of infants observed to have tested positive fell sharply over time from nearly 7% among those born in 2010 to 1% among those born in 2016. It is important to note that these estimates may not reflect the true MTCT rate, as some HIV-exposed infants will have never received a PCR test and many intrapartum and post-partum infections are likely to have been missed due to a lack of repeat testing after birth. It is also possible that some HIV-unexposed infants were also tested, leading to an underestimation of the true rate; in MONARCH, 2% of infants not exposed to HIV received a PCR test (see Chapter 5). In addition, since data from rapid antibody tests are not systematically collected anywhere, only infants acquiring HIV before 18 months of age could be captured. The median age at HIV diagnosis increased after compared to before the introduction of birth testing, although the difference was not statistically significant.

This analysis has several other limitations not described above. Firstly, only the receipt of PCR testing was considered here, as no data were available to assess the rest of the early infant diagnosis pathway, for example how many results were returned to first the clinic and then to caregivers, and how long this took. A study assessing return of results at a single hospital in Johannesburg between 2014 and 2016 reported that the results from only 52% of birth tests were returned to mothers, in a median (IQR) time of 10 (9, 13) days [155]. Secondly, missed matches in the within NHLS linkage would have resulted in an underestimation of the proportion repeat tested, though in turn this would mean overestimation of the proportion of infants ever tested as estimated in Chapter 5. However, given the number of tests conducted at each age, it is clear that under the new guidelines fewer infants were tested at 10-18 weeks than at birth, for example. Thirdly, residents within the sub-district are a highly mobile population, with migration events from inside to outside the surveillance area estimated at 61 per 1,000 person-years among infants (personal communication: K. Herbst, 2018). This could mean that some infants may have had repeat tests conducted at clinics outside of the sub-district which would not have been captured here. Equally, infants who appear to not have been tested at birth may in fact have received a birth test elsewhere. Finally, since data from one of the clinics in the sub-district were only available between January 2011 and September 2015, repeat tests on some infants may have been missed in the later part of the time period.

7.6. Key findings

The key findings from this chapter are:

- Among those born after the introduction of birth testing, less than a quarter of infants who tested negative at birth ever received another test, which may result in undiagnosed intrapartum and post partum infections

- Repeat testing after an initial positive result was also low, which may result in infants who do not have HIV (but receive a false positive result) being initiated on treatment

- Caregiver-reported vaccination coverage was high suggesting good engagement in care of infants; this may mean that poor testing coverage may be due to inadequate linkage across maternal and child services and follow-up procedures in clinic

- Adherence to PCR testing guidelines varied by clinic, which may provide an opportunity to identify best practices in the better performing clinics, to optimise service delivery across all clinics in this setting

Chapter 8. ART initiation and viral suppression

8.1. Introduction

In Chapter 7, I estimated the proportion of HIV-exposed infants who received a positive PCR test result, and in this chapter I describe the proportion of these infants who went on to start ART, based on the linkage between NHLS and ACCDB. I subsequently describe their characteristics at ART initiation and outcomes on treatment, including viral suppression.

As data for outcomes on ART extend beyond 2 years of age the term 'children' is used throughout this chapter rather than 'infants', however all analyses are still based on the same group of individuals born in the sub-district between 1st June 2010 and 31st December 2016.

8.2. Objectives

The objectives for this chapter are to:

- Estimate the proportion of children diagnosed with HIV who went on to initiate ART
- Among those on ART, describe characteristics at ART initiation including age, initial regimen, and immunological stage
- Describe outcomes after ART initiation, including the frequency of viral load and CD4 monitoring, viral suppression, treatment interruptions and changes to treatment regimens, and mortality

8.3. Methods

8.3.1. ART coverage

ART coverage among those with a linked positive PCR test

Among those with a positive PCR test in the NHLS dataset, the proportion of children ever initiating ART (defined as being linked to ACCDB, the AHRI ART database, through linkage conducted in Chapter 4) was summarised, overall and by year of birth. Among those who initiated ART within the sub-district and therefore had a known date of initiation (as opposed to having initiated ART at a clinic elsewhere and later transferring in to the sub-district) for whom data prior to transfer, including the date of their initiation, age at initiation and time to initiation from first positive PCR test result were summarised. The change over calendar time in the time from first positive result to ART initiation was assessed using a Cox model, using date of birth as a linear continuous predictor variable, with the time to initiation set to zero for any individuals who initiated on ART on the same day or prior to the date of their diagnosis. The proportion

initiating ART by 6 months after first positive result was summarised, with those diagnosed after the 31st December 2016 (in whom ART data were not available up to 6 months of age) excluded.

Between 2010 and March 2013, the South African guidelines recommended immediate initiation of ART for those <1 year of age but delayed initiation for children aged >1 year until their CD4 fell below 750 cells/mm^3 or 25%, or they reached WHO clinical stage 3/4. To explore the impact of this on results, age at and time to ART initiation were additionally summarised excluding those diagnosed before March 2013 at >1 year of age.

In order to explore the impact of migration on estimated ART coverage, the proportion of diagnosed children ever starting ART was summarised among those who were linked to ACDIS and were known to still be resident in the demographic surveillance area for at least 1 year after diagnosis, and was compared to the proportion of those who were not still resident using a chi-squared test.

Among those linked to ACDIS, the follow-up status of those with a positive PCR test but who never started ART was described; the proportion known to have died was reported, as well as the age at, cause of and location of death where available.

Finally, ART coverage among those born after the introduction of birth testing on 1st April 2015 was summarised in more detail. Since this time, national guidelines have recommended the immediate initiation of treatment upon receipt of a positive result, alongside a confirmatory PCR test (rather than delaying initiation until the result of the confirmatory test was received). The timing of ART initiation after an initial positive test relative to subsequent PCR testing was therefore summarised. In addition, ART coverage and timing of ART initiation was compared between those diagnosed at birth (defined as <7 days of age) and those diagnosed at an older age, using a chi-squared test and Wilcoxon's rank-sum test.

Characteristics of children on ART but with no linked positive PCR test

The characteristics of children born between 1st July 2010 and 31st December 2016 who were reported to have ever been on ART at one of the healthcare facilities in the sub-district but who were not linked to any positive PCR tests were summarised, including the median age at last recorded negative or indeterminate PCR test (among those who had one) and the median age first seen on ART within the sub-district.

The proportion of these individuals who may not be expected to have had a positive PCR test within the sub-district was summarised. There were two possible reasons for this. Firstly, some initiated ART outside the sub-district and were therefore likely to have also been diagnosed outside of the sub-district. Secondly, some initiated ART >18 months of age and were therefore likely to have been diagnosed with an antibody test, rather than a PCR test. The remaining

children were those who were likely to have received a positive PCR test within the sub-district but for whom the test result was not linked, termed 'potential missed links'.

In order to assess the possible underestimation of ART coverage among those diagnosed with HIV as a result of these potential missed links, a sensitivity analysis was conducted. For this, those who had a positive PCR test but who did not start ART were assumed to overlap with the group who were on ART with no positive PCR test; based on this, an upper bound of the estimate of ART coverage was calculated.

8.3.2. Characteristics at ART initiation

For those who initiated ART in the sub-district (and thus had data available from initiation), characteristics at ART initiation were summarised. Age at initiation was described, and initial regimens were summarised, overall and by age at initiation and year of initiation. Age categories were chosen to reflect guidelines and recommendations for the use of specific drugs in each ART drug class, with details in Table 8.1; the use of each third agent (PI or NNRTI) was summarised separately for those <14 days, 14 to <28 days, 28 days to <3 months, 3 months to <3 years and ≥3 years, and each NRTI backbone for those <28 days, 28 days to <3 months, 3 months to <2 years and ≥2 years. In order to assess the quality of recording of ART data in TIER.net (on which most ART data in ACCDB is based), initial regimens among children born after July 2010 were also summarised using data from ARTemis, the previous AHRI clinical database. Among those who were linked between the two datasets, initial regimens recorded in each were compared at an individual level.

Table 8.1 - Reasons for the choice of cut-offs for age categorisations of initial third agent (PI or NNRTI) and NRTI backbone.

Third agent (PI or NNRTI)		NRTI backbone	
14 days	LPV/r not recommended for neonates under 14 days	28 days	South African guidelines recommend specialist supervision when choosing regimens for those <28 days
28 days	South African guidelines recommend specialist supervision when choosing regimens for those <28 days		
3 months	EFV sprinkles licensed by FDA for those >3 months	3 months	ABC not normally recommended for those <3 months
3 years	EFV tablets not normally recommended until 3 years of age	2 years	TDF licensed by FDA for those >2 years

Finally, the immunological status of children was summarised using the CDC 2014 surveillance classification [80], based on the CD4 measurement closest to ART initiation within a window of

6 months before and 1 month after ART initiation. CDC staging was chosen as this has criteria based on both CD4 counts and CD4% for younger children, rather than just CD4% as the WHO classification does, meaning immunological stage could be summarised for more children. For children who had both a CD4 count and CD4% taken on the same day which would classify them in different categories, the most severe category was used. CDC stage was summarised by calendar year of initiation and by age.

8.3.3. Outcomes after ART initiation

Outcomes on ART were summarised, in some cases just restricted to those with data available from initiation, and in others also including those who transferred in to the sub-district on ART (i.e. for whom the date of initiation and ART data prior to transfer was not available).

The most recent follow-up status of each child was categorised as either still in follow-up, transferred clinic, lost to follow-up, or died. Children were considered lost-to-follow-up if they were reported as such by their clinic and were not known to be on ART at another clinic, or if they had no ART data recorded in ACCDB during the 6 months prior to data extraction for the most recent version. Children were known to have died either because this had been recorded in ACCDB, or because it had been reported through demographic surveillance, for children who had been linked to ACDIS. Those still in follow-up were censored at the date of their last visit.

Among children with >1 day follow-up on ART, the rates of CD4 and viral load monitoring were summarised. The rates were calculated as the number of viral load/CD4 tests conducted divided by the total duration of follow-up. These rates were calculated overall, by clinic, by number of years on ART (among those with data from ART initiation only), and by calendar year. The correlation between the rates of viral load and CD4 monitoring at each clinic was calculated, as was the correlation between the size of each clinic (measured using the total duration of follow-up) and each of the rate of viral load and CD4 testing. In order to explore the completeness of recording of viral load and CD4 data in TIER.net (and thus in ACCDB), a comparison was made to ARTemis; among children linked between the two datasets, the proportion with a measurement prior to 20th January 2014 (the data cut-off for ARTemis) recorded in ARTemis but not in ACCDB (and vice versa) was summarised.

Viral suppression was defined as a viral load ≤400 copies/mL. For those with data from ART initiation, the time to and the proportion suppressed at first viral load measurement was summarised, along with the proportion suppressed at 12 months after initiation (allowing for a window of ±3 months). Time to suppression from ART initiation was estimated using Kaplan-Meier methods (among those with data from initiation only), censoring children at their last follow-up date. Among all children on ART, the proportion suppressed at last measurement was summarised, overall and by current age (categorised at 0-<1, 1-<3, 3-<5 and ≥5 years) and

regimen (LPV/r- vs EFV-based). An upper and lower bound on the proportion suppressed at their most recent measurement was calculated assuming those with no viral load data were all suppressed and unsuppressed respectively. Time to the next viral load measurement following a viral load >1,000 copies/mL was summarised.

CDC immunological stage at most recent CD4 measurement was described, as well as the proportion at a more or less severe stage compared to at ART initiation. Changes in actual CD4 counts and percentages were not considered due to the difficulty in interpretation caused by age-related changes to children's immune systems.

In TIER.net, the number of days of ART prescribed at each visit is recorded by the clinic. From this, the expected date of the next visit is calculated in ACCDB by AHRI. A treatment interruption was defined as any gap of ≥1 day between the expected next visit date and the actual date of next visit. The number of individuals with a treatment interruption was summarised, as well as the length of their longest interruption. The medication possession ratio was calculated as the number of days with ART, divided by the total duration of follow-up since ART initiation [252]. The medication possession ratio was additionally summarised by clinic, with time off ART that preceded a change of clinic excluded. The median duration of ART prescriptions given was calculated based on the time between visits when no interruptions occurred. In order to better convey the frequency, timing and duration of interruptions, treatment data were plotted for a random sample (15%, for readability) of children, with time on and off ART indicated.

Treatment changes were explored among those with data from ART initiation and with a known initial ART regimen. Four types of treatment change were considered:

i)   any change, defined as a change to any drug

ii)   a change to the third agent, defined as any change to either the NNRTI or PI, including within class changes

iii)   a change to the NRTI, defined as any change to the NRTI, excluding changes between 3TC and FTC, which are considered therapeutically interchangeable [253]

iv)   a switch to second-line, defined as a simultaneous change to the NRTI backbone and a third agent. The number and proportion of children who ever experienced each type of change was summarised. To investigate the impact of potential data entry errors on the results, a sensitivity analysis was conducted ignoring changes to treatment where the new regimen was only recorded at one visit, and the regimens recorded at the visits immediately before and after were the same.

Switches to second-line were described in more detail. The cumulative incidence of switch to second-line was summarised using Kaplan-Meier, overall, by age at initiation (<28 days, 28 days to <3 months, 3 months to <2 years, and ≥2 years), by year of ART initiation and by initial regimen

(LPV/r-based vs. EFV-based). Reasons for treatment changes are not collected in TIER.net (and thus are not available in ACCDB), but in order to explore this, CD4 and viral load measurements available in the 3 months prior to switch were summarised, as well as the proportion of switches that occurred after a break from treatment, and the proportion that occurred following a change of clinic.

8.4. Results

8.4.1. ART coverage

<u>ART coverage among those with a linked positive PCR test</u>

Of 458 children who had a positive PCR test, 244 (53%) were identified as having subsequently initiated ART through linkage to the ART database (Table 8.2 and Figure 8.1). There was evidence of a change in the proportion initiating over time (p=0.006); the proportion increased from 45% among those born in 2010 to 70% among those born in 2015, although there was a subsequent decrease to 58% among those born in 2016.

Table 8.2 - Number and proportion of children with a positive PCR test ever initiating ART, by year of birth

Year of birth	Number (%) initiating ART
2010	31/69 (45%)
2011	50/96 (52%)
2012	40/78 (51%)
2013	30/67 (45%)
2014	34/60 (57%)
2015	45/64 (70%)
2016	14/24 (58%)
Overall	244/458 (53%)
p	0.006

Note: p-value for change in proportion over time

Figure 8.1 - Proportion of children with a positive PCR test ever initiating ART, by year of birth
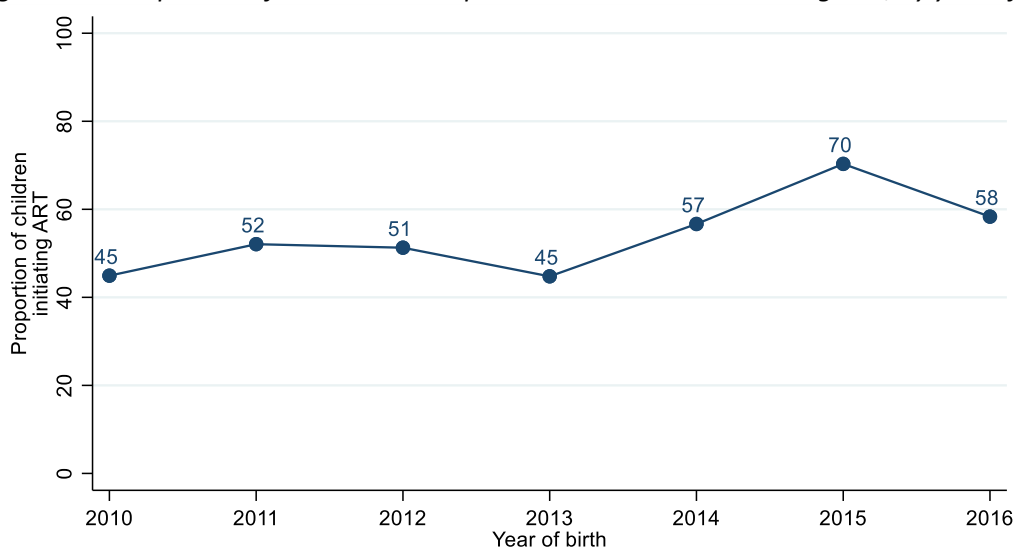
The date of ART initiation was available for 205/244 children; 38 had transferred to a clinic within the sub-district after initiating ART elsewhere so the date of their initiation was unknown (as data prior to transfer was not available through ACCDB) and one was reported in ACCDB to have started treatment but had no ART follow-up data recorded.

Among those children with a known date of ART initiation, the median (IQR) time from first positive test to ART initiation was 5.0 (1.9, 12.4) weeks, with 26 (13%) and 10 (5%) not starting until 6 months and 1 year after their first positive test respectively (Table 8.3 and Figure 8.2). Four children (2%) started ART before the date of their first positive PCR test (1 started 7 days before at age 2 weeks, and 3 started 2 days before at ages 20, 31 and 33 weeks) and 3 (1%) started on the same day as their first positive test. The median time between first positive test and ART initiation remained relatively stable among those born between 2010 and 2013 at between 7.1 and 8.1 weeks, after which it decreased steadily to 1 week among those born in 2016 (p<0.001).

Table 8.3 - Time from first positive PCR test result to ART initiation, among all children and excluding those diagnosed before March 2013 aged >1 year, by year of birth

Year of birth	Time from first positive PCR test result to ART initiation (weeks)		Time from first positive PCR test result to ART initiation, excluding those diagnosed before March 2013 aged >1 year (weeks)	
	n	median (IQR) [range]	n	median (IQR) [range]
2010	25	8.0 (4.7, 18.9) [1.4, 73.9]	22	7.1 (3.9, 18.9) [1.4, 73.9]
2011	41	8.9 (5.1, 22.9) [0.4, 175.4]	40	9.2 (5.1, 23.0) [0.4, 175.4]
2012	32	7.1 (3.7, 18.0) [1.3, 140.9]	32	7.1 (3.7, 18.0) [1.3, 140.9]
2013	24	8.1 (3.2, 19.6) [0.7, 103.7]	24	8.1 (3.2, 19.6) [0.7, 103.7]
2014	29	4.4 (2.6, 7.3) [0.3, 48.9]	29	4.4 (2.6, 7.3) [0.3, 48.9]
2015	43	1.7 (1.1, 2.4) [-1.0, 22.1]	43	1.7 (1.1, 2.4) [-1.0, 22.1]
2016	11	1.0 (0.6, 2.9) [-0.3, 37.0]	11	1.0 (0.6, 2.9) [-0.3, 37.0]
Overall	205	5.0 (1.9, 12.4) [-1.0, 175.4]	201	4.9 (1.9, 12.4) [-1.0, 175.4]
p		<0.001		<0.001

Note: p-value for change in time from first positive PCR test result to ART initiation over time

Figure 8.2 - Time from first positive PCR test result to ART initiation, among all children and excluding those diagnosed before March 2013 aged >1 year, by year of birth
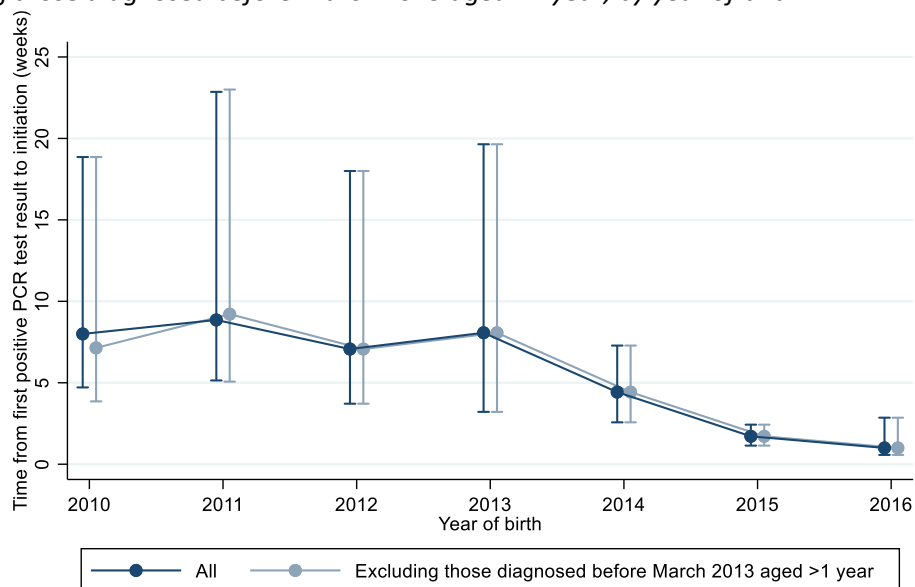
The time from first positive PCR test result to ART initiation is shown in Figure 8.3, by year of birth, excluding those diagnosed after 31st December 2016. Those born in more recent calendar years were more likely to start treatment earlier, with the proportion (95% CI) initiating by 6 months increasing from 33.3% (23.6%, 45.8%) among those born in 2010 to 68.9% (57.1%, 80.0%) in 2015 (Table 8.4). The proportion among those initiating in 2016 was lower, though with a wider confidence interval given the lower number of children in this group. The hazard ratio (HR) per year increase in date of birth was estimated at 1.18 (95% CI 1.10, 1.27), p<0.001.

Table 8.4 - Proportion of children initiating ART by 6 months of age, by year of birth

Year of birth	Proportion of children initiating ART by 6 months of age (95% CI)
2010	33.3% (23.6%, 45.8%)
2011	40.0% (30.6%, 50.1%)
2012	41.0% (31.1%, 52.8%)
2013	34.3% (24.3%, 47.0%)
2014	50.0% (38.2%, 63.2%)
2015	68.9% (57.1%, 80.0%)
2016	47.1% (27.0%, 72.4%)
Overall	43.8% (39.3%, 48.5%)

7 children diagnosed after 31st December 2016 were excluded

Figure 8.3 - Time to ART initiation from first positive PCR test result, by year of birth
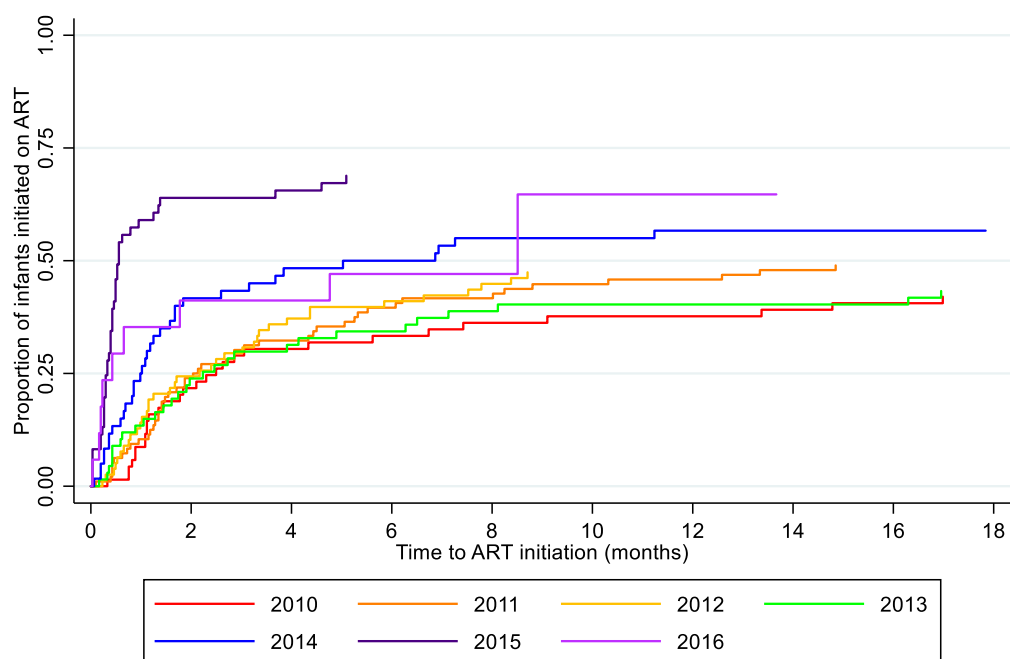
The median (IQR) age at ART initiation was 25.7 (14.1, 44.7) weeks (Table 8.5 and Figure 8.4), decreasing over time from 32.1 (16.6, 50.4) weeks among those born in 2010 to 20.7 (13.3, 37.6) weeks among those born in 2016 (p<0.001). There was little change in either the median time to initiation or the median age at initiation after excluding 4 children who were diagnosed before March 2013 when they were older than 1 year of age.

Table 8.5 - Age at ART initiation, among all children and excluding those diagnosed before March 2013 aged >1 year, by year of birth

Year of birth	Age at ART initiation (weeks)		Age at ART initiation, excluding those diagnosed before March 2013 aged >1 year (weeks)	
	n	median (IQR) [range]	n	median (IQR) [range]
2010	25	32.1 (16.6, 50.4) [11.7, 90.9]	22	28.3 [15.7, 48.1] [11.7, 80.3]
2011	41	31.7 (19.4, 52.4) [10.9, 183.1]	40	30.5 (19.1, 52.0) [10.9, 183.1]
2012	32	22.7 (13.1, 36.6) [7.1, 148.0]	32	22.7 (13.1, 36.6) [7.1, 148.0]
2013	24	35.1 (16.1, 51.2) [7.1, 129.9]	24	35.1 (16.1, 51.2) [7.1, 129.9]
2014	29	27.0 (13.3, 40.9) [8.7, 88.9]	29	27.0 (13.3, 40.9) [8.7, 88.9]
2015	43	19.7 (10.3, 31.0) [1.3, 76.9]	43	19.7 (10.3, 31.0) [1.3, 76.9]
2016	11	20.7 (13.3, 37.6) [1.0, 64.1]	11	20.7 (13.3, 37.6) [1.0, 64.1]
Overall	205	25.7 (14.1, 44.7) [1.0, 183.1]	201	25.4 (14.0, 43.7) [1.0, 183.1]
p		<0.001		0.039

Note: p-value for change in age at initiation over time

Figure 8.4 - Age at ART initiation, among all children and excluding those diagnosed before March 2013 aged >1 year, by year of birth
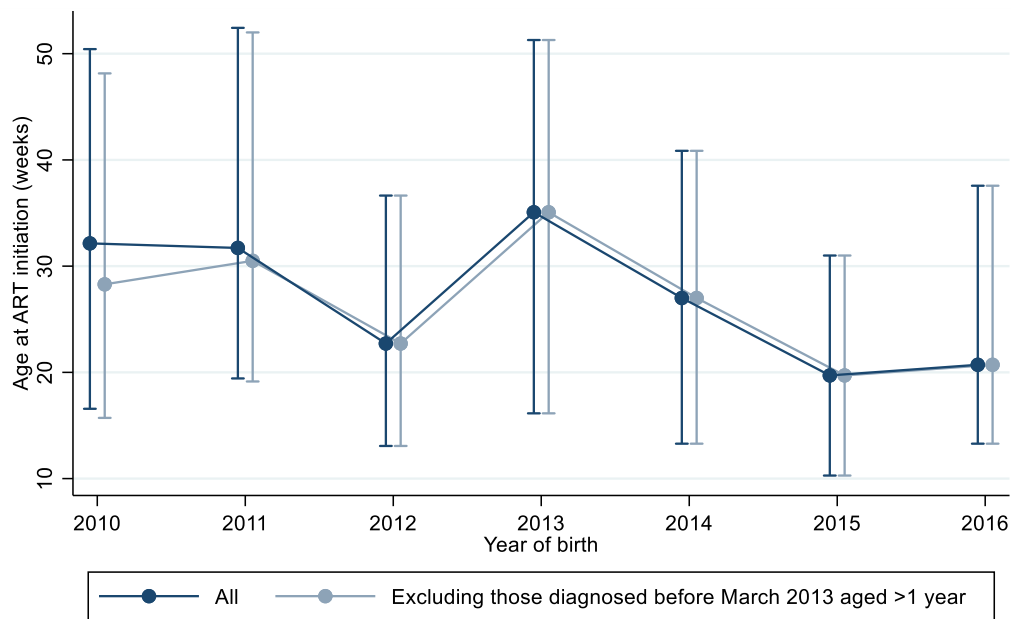

Among 73 children who were linked to ACDIS and had residency data, 40 were still resident in the demographic surveillance area 1 year after first testing positive and 33 were not. There was no difference in the proportion who initiated ART between these two groups (28/40 (70%) among those still resident compared to 22/33 (67%) of those not, p=0.760).

Among 23 children who were linked to ACDIS, and who tested positive but never initiated ART, 10 (43%) were known to have died, with a resultant mortality rate of 15.0 (95% CI 8.1, 27.8) per 100 person-years. Among those who died, the median (IQR) [range] age at death was 6.5 (3.2, 7.7) [1.8, 16.0] months, and the time from diagnosis to death was 2.9 (0.7, 4.8) [0.3, 14.3] months. The location of death was home for 5 (50%), a healthcare facility for 4 (40%), and was unknown for 1 (10%). For 7 children the cause of death was recorded; 3 (43%) died from an acute respiratory infection including pneumonia, 1 (14%) from diarrhoeal diseases, and the cause was only coded as HIV/AIDS related for the remaining 3 (43%). There was no change in the proportion who died by year of birth (p=0.322).

Among 68 children born after the change to the introduction of birth testing into the national testing guidelines on 1st April 2015 and who had a positive PCR test, 13 were diagnosed at birth and 55 were diagnosed at an older age. A comparison of ART coverage between these two groups is shown in Table 8.6. A lower proportion of those diagnosed at birth went on to initiate ART (46% vs. 73%, p=0.065), although the median age at initiation in this group was substantially younger (1.4 vs. 19.7 weeks, p=0.020). Of the 46 children who went on to initiate ART, 17 (37%) initiated ART at the same time as or prior to any subsequent PCR tests, 10 (22%) initiated ART after at least one confirmatory PCR test, and the remaining 19 (41%) had no subsequent PCR tests after their initial positive result.

Table 8.6 - ART coverage among those diagnosed at birth compared to those diagnosed later, for those born after the change to the national testing guidelines on 1st April 2015

	Diagnosed at birth	Diagnosed after birth	Total	p
	n (%) or median (IQR)			
Number testing positive	13	55	68	
Proportion initiating ART	6 (46%)	40 (73%)	46 (68%)	0.065
Number with known date of ART initiation	5	37	42	
Time from diagnosis to initiation (weeks)	1.4 (1.4, 2.1)	1.7 (1.0, 2.3)	1.6 (1.0, 2.3)	0.627
Age at initiation (weeks)	1.4 (1.3, 2.1)	19.7 (12.4, 31.4)	18.0 (10.3, 31.4)	0.020

Note: Diagnosis at birth defined as first positive test <7 days of age

<u>Children on ART but with no linked positive PCR tests</u>

In addition to the 244 children with a positive PCR test result who were on ART at one of the healthcare facilities within the sub-district, there were 483 children born between 1st June 2010 and 31st December 2016 who were reported to be on ART in ACCDB but who did not have any linked positive PCR test results. A flowchart describing the availability of ART data, the place of ART initiation and PCR test data among these children is shown in Figure 8.5.

Of these 483 children, 48 were reported to have initiated ART on the day that they were born. By comparison, none of the children with a linked positive PCR test result started ART on the day they were born. The reported regimens were all triple therapy, rather than regimens recommended in the national guidelines for the prevention of mother to child transmission. Given the turnaround time required to process PCR tests and return the results to the patient's caregivers in this setting, it was thought to be very unlikely that a child would be able to start treatment on the day they were born. Six of these individuals had been linked to ACDIS, in which all had a different date of birth reported, with the year of birth recorded in ACDIS ranging from 1966 to 1995. It was therefore assumed that for all 48 individuals the date of ART initiation had been mistakenly also entered for the date of birth, and they were therefore removed from all subsequent analyses.

Among the remaining 435 children, 396 (91%) had no linked PCR tests and 39 (9%) children had at least one linked negative or indeterminate PCR test, of whom 29 had only negative test results and 10 had indeterminate test results (including one who had also previously had a negative result, but whose most recent test result was indeterminate). For each of these three groups, a summary of the age first seen on ART within the sub-district, age at last linked PCR test, and gap between last linked PCR test and first time seen on ART (where applicable) is shown in Table 8.7, separately for those who initiated ART in the sub-district compared to those to initiated elsewhere and later transferred in.

Figure 8.5 - Flowchart of children reported to be on ART in ACCDB but with no linked positive PCR test results
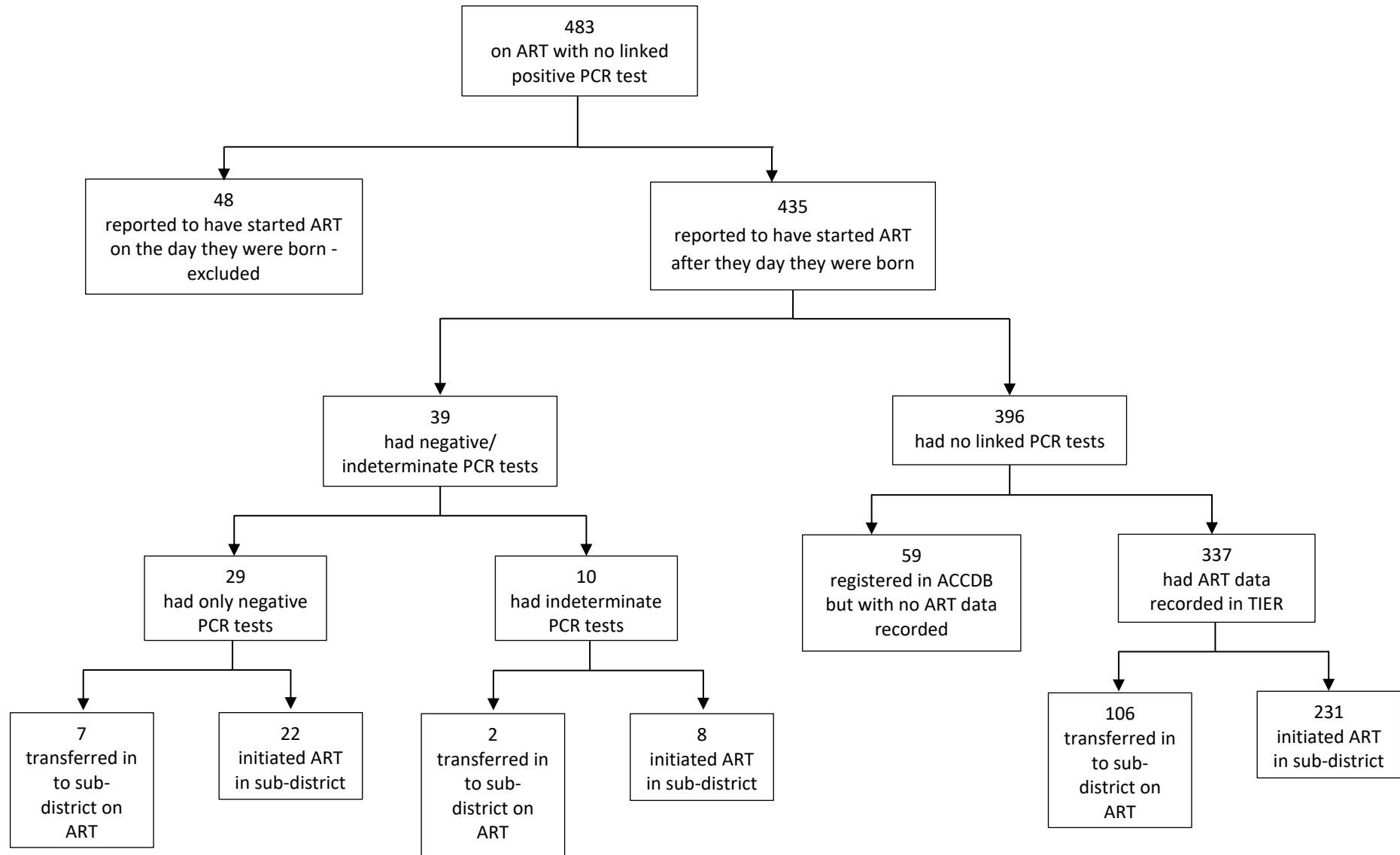
Table 8.7 - Age at last linked PCR test, age first seen on ART within the sub-district, and gap between last linked PCR test and first time seen on ART, among those on ART but with no linked positive PCR test result

		n	Age at last PCR test (weeks)	Age first seen on ART (weeks)	Gap between last PCR test and first seen on ART (weeks)
			median (IQR) [range]		
Negative only (N=29)	ART initiation	22	6.6 (6.1, 9.3) [0.0, 40.6]	90.6 (80.0, 144.6) [40.6, 212.4]	84.6 (73.1, 113.7) [0.0, 201.6]
	Transferred in	7	10.0 (6.1, 24.4) [6.1, 49.0]	83.4 (81.1, 131.1) [78.1, 165.3]	73.4 (57.9, 125.0) [40.1, 158.0]
Indeterminate ever (N=10)	ART initiation	8	56.4 (24.1, 68.6) [10.0, 82.4]	73.6 (56.1, 95.9) [46.3, 136.1]	17.1 (-0.4, 42.6) [-1.7, 126.1]
	Transferred in	2	9.8 (1.7, 17.9) [1.7, 17.9]	21.8 (17.1, 26.4) [17.1, 26.4]	12.0 (8.6, 15.4) [8.6, 15.4]
No linked PCR tests (N=337)	ART initiation	231	-	82.0 (26.1, 128.6) [0.3, 339.9]	-
	Transferred in	106	-	102.1 (61.6, 155.9) [4.7, 318.3]	-

Note: Age first seen on ART represents age initiated ART for those who initiated ART within the sub-district, and age transferred in to the sub-district for those who initiated elsewhere

Among the 29 children with only negative test results, 7 (24%) initiated ART outside the sub-district. They were last tested at a median (IQR) 10.0 (6.1, 24.4) weeks of age and then transferred in to a sub-district clinic on ART at a median of 73.3 (57.9, 125.0) weeks later. The remaining 22 (76%) were last tested at a median 6.6 (6.1, 9.3) weeks of age, and subsequently initiated ART within the sub-district at a median 84.6 (73.1, 113.7) weeks later, with 19 (95%) and 15 (75%) initiating after the age of 12 and 18 months of age respectively.

Among the 10 children with an indeterminate PCR test result, 2 (20%), who were last tested at 17.9 and 1.7 weeks of age, subsequently initiated ART outside the sub-district and later transferred to a facility within the sub-district 8.6 and 15.4 weeks later respectively. The remaining 8 (80%), who were last tested at 56.4 (24.1, 68.6) weeks of age (all indeterminate), initiated ART at a median (IQR) age of 17.1 (-0.4, 42.6) weeks, with 7 (88%) and 4 (50%) initiating after the age of 12 and 18 months of age respectively.

Of the 396 children with no linked PCR tests, 59 were registered in ACCDB indicating that they had initiated treatment, but had no ART data recorded. Of the remaining 337, 106 (31%) initiated ART outside of sub-district and transferred in at a median (IQR) age of 102.1 (61.6, 155.9) weeks, and 231 (69%) initiated ART within sub-district at a median (IQR) age of 82.0 (26.1, 128.6) weeks of age, with 152 (66%) and 131 (57%) initiating after the age of 12 and 18 months of age respectively.

A summary of all those on ART with no linked positive PCR test result is shown in Table 8.8. Excluding the 59 who were registered in ACCDB but had no ART data, 376 children with available

ART data remained. Of these, 115 (31%) initiated ART outside of the sub-district (and therefore were likely to have also been diagnosed outside of the sub-district) and 152 (40%) initiated ART after the age of 18 months (and therefore may have been diagnosed using an antibody test rather than a PCR test). The remaining 109 (29%) represent potential missed links between NHLS PCR test data and ACCDB, that is those who would be expected to have received a positive PCR test result at a healthcare facility within the sub-district but for whom one was not linked.

Table 8.8 - PCR testing status and characteristics at ART initiation of children reported to be on ART within the sub-district but with no linked positive PCR test

| | No linked PCR test | Had linked PCR test | | Total |
		All negative	Ever indeterminate	
Total	396	29	10	435
No ART data recorded in ACCDB	59	0	0	59
ART data recorded in ACCDB	337	29	10	376
Initiated ART outside of the sub-district	106 (31%)	7 (24%)	2 (20%)	115 (31%)
Initiated ART >18 months of age	131 (39%)	17 (59%)	4 (40%)	152 (40%)
Possible missed NHLS-ACCDB links	100 (30%)	5 (17%)	4 (40%)	109 (29%)
Initiated ART between 12-18 months	21	4	3	28

As a sensitivity analysis, the impact of these 109 potential missed links on the estimate of ART coverage among those with a positive PCR test result was assessed. It was assumed that these explained a proportion of the group of 214 children with a positive PCR test result who did not go on to initiate ART. In this case, the proportion not on ART would be (214-109)/458 = 105/458 = 23%, with the corresponding upper bound on the estimated ART coverage being 77%, compared to the original estimate of 53%.

8.4.2. Characteristics at ART initiation

In total, there were 679 children born between 1st June 2010 and 31st December 2016 who were reported to be on ART in ACCDB (including both those with and without linked positive PCR tests), of whom 60 (9%) had no data on ART recorded (Table 8.9). Of the remaining 619 (91%) with ART data available, 466 (75%) initiated ART within the sub-district and 153 (25%) initiated ART elsewhere and subsequently transferred to a clinic within the sub-district.

Table 8.9 - PCR testing status and ART initiation characteristics of all children on ART in the sub-district

	Linked positive PCR test	No linked positive PCR test	Total
Total	244	435	679
No ART data recorded in ACCDB	1	59	60
ART data recorded in ACCDB	243	376	619
Initiated ART in sub-district	205 (84%)	261 (69%)	466 (75%)
Initiated ART elsewhere and later transferred in	38 (16%)	115 (31%)	153 (25%)

Age at ART initiation

For the 466 children with data available from ART initiation, age at ART initiation is shown in Figure 8.6. The median (IQR) [range] age at initiation was 10.7 (4.2, 21.0) months [2 days, 6.5 years], with 17 (4%) starting <28 days of age, of whom 4 and 10 started <7 and <14 days respectively, and 97 (21%) starting ≥2 years of age.

Figure 8.6 - Age at ART initiation among those with data available from ART initiation
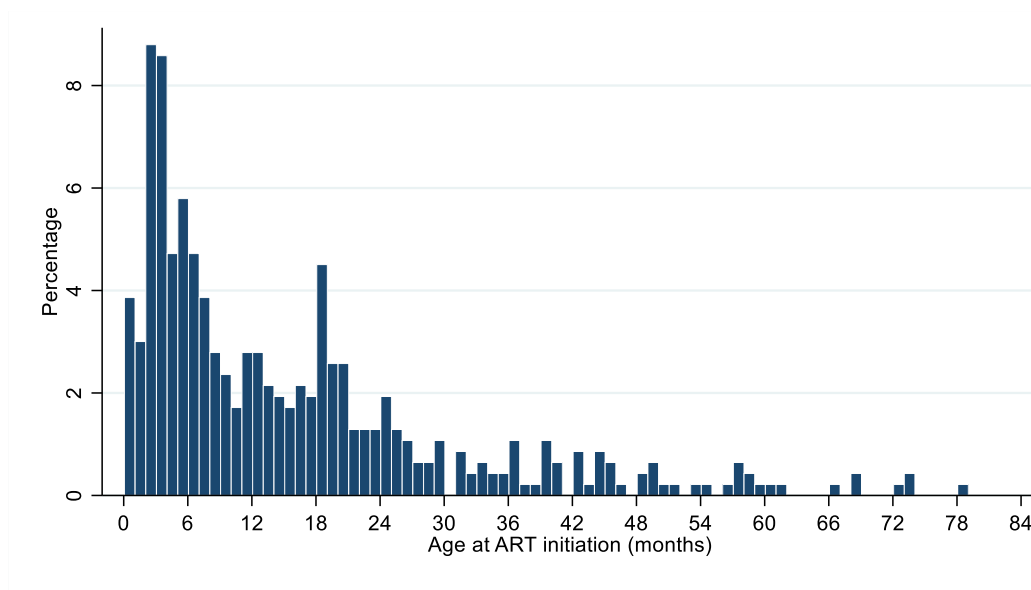
Initial regimen

A summary of the initial regimens used is shown in Table 8.10; 308 (66%) children initiated ART on a LPV/r-based regimen, 136 (29%) on an EFV-based regimen, 3 (1%) on an NVP-based regimen, 3 (1%) on another regimen, and the initial regimen was unknown for the remaining 16 (3%).

Table 8.10 - Initial ART regimens among those with data available from ART initiation

Third agent	Regimen	n (%)
LPV/r + 2NRTI (n=308, 66%)	LPV/r + ABC + 3TC	297 (64%)
	LPV/r + ZDV + 3TC	8 (2%)
	LPV/r + d4T + 3TC	2 (<0.5%)
	LPV/r + TDF + 3TC	1 (<0.5%)
EFV + 2NRTI (n=136, 29%)	EFV + ABC + 3TC	81 (17%)
	EFV + TDF + FTC	38 (8%)
	EFV + TDF + 3TC	10 (2%)
	EFV + d4T + 3TC	6 (1%)
	EFV + ZDV + 3TC	1 (<0.5%)
NVP + 2NRTI (n=3, 1%)	NVP + ZDV + 3TC	2 (<0.5%)
	NVP + TDF + FTC	1 (<0.5%)
Other (n=3, 1%)	RTV + ABC + 3TC	2 (<0.5%)
	ABC + 3TC	1 (<0.5%)
Unknown (n=16, 3%)	-	

A summary of the third agent used, by age at initiation, is shown in Table 8.11, and by age and year of ART initiation in Figure 8.7. The majority of those who initiated ART <14 days of age started on an EFV-based (5/10, 50%) or LPV/r-based (4/10, 40%) regimen. Most children aged 14 days to 3 years initiated on a LPV/r-based regimen, although 21/57 (37%) of those initiating 28 days to <3 months of age and 73/328 (22%) initiating 3 months to <3 years started on an EFV-based regimen. Among those initiating ART after 3 years of age, the majority (70%) started on an EFV-based regimen. Only 3 (1%) children (all aged <3 months) initiated on a NVP-based regimen. There was no clear trend in third agent use over calendar time.

Table 8.11 - Initial third agent, by age at ART initiation, among those with data available from ART initiation

Age at ART initiation	LPV/r	EFV	NVP	None	Unknown	Total
<14 days	4 (40%)	5 (50%)	1 (10%)	0	0	10
14 - <28 days	4 (57%)	2 (29%)	1 (14%)	0	0	7
28 days - <3 months	29 (59%)	19 (39%)	1 (2%)	1	3	53
3 months - <3 years	255 (78%)	73 (22%)	0	2	13	343
≥3 years	16 (30%)	37 (70%)	0	0	0	53
Total	308 (69%)	136 (30%)	3 (1%)	3	16	466

Note: Percentages exclude those with no/unknown third agent.

Figure 8.7 - Initial third agent, by age at ART initiation and year of ART initiation, among those with data available from ART initiation
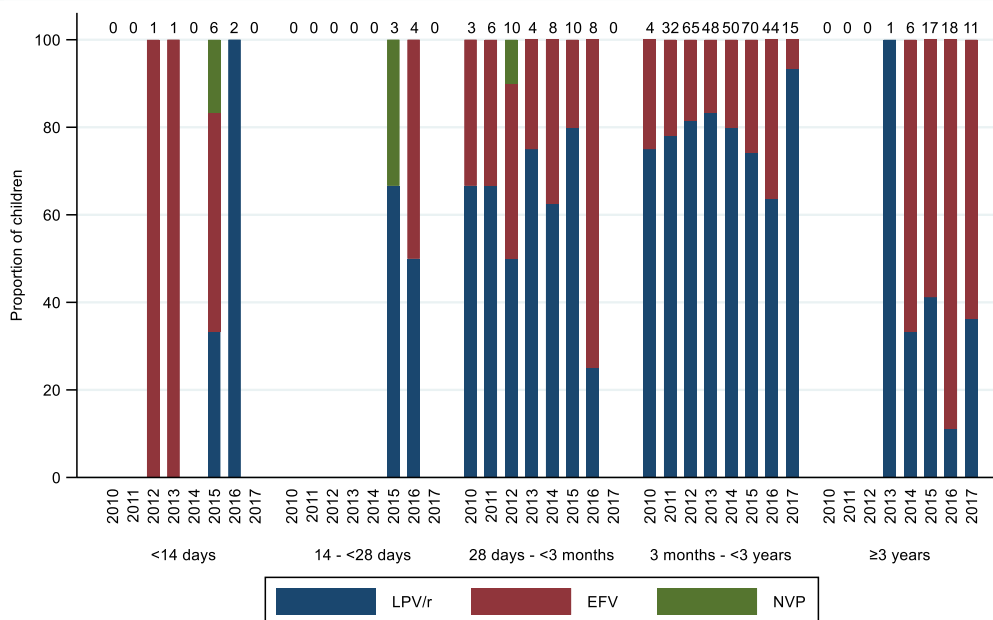
Notes: The number above each bar represents the number of children born in each year initiating ART in each age group. Figure includes 447 children with initial regimen of form 2NRTI + 1NNRTI/PI.

A summary of the NRTI backbone used, by age at initiation, is shown in Table 8.12, and by age and year of ART initiation in Figure 8.8. ABC+3TC was the most common backbone used among all age groups, including those aged <28 days (59%, 10/17), with the proportion using this

combination increasing with age. Overall, 11% (50/450) of children were reported to have started treatment with TDF+3/FTC, including 24% (4/17), 28% (14/53) and 10% (28/299) of those aged <28 days, 28 days to <3 months and 3 months to <2 years respectively. A small number of children were initiated on ZDV+3TC, almost all of whom started treatment in 2015, and on d4T+3TC, mostly recently in 2015.

Table 8.12 - Initial NRTI backbone, by age at ART initiation, among those with data available from ART initiation

Age at ART initiation	ABC + 3TC	ZDV + 3TC	TDF + 3/FTC	d4T + 3TC	Unknown	Total
<28 days	10 (59%)	2 (12%)	4 (24%)	1 (6%)	0	17
28 days - <3 months	35 (70%)	1 (2%)	14 (28%)	0	3	53
3 months - <2 years	250 (87%)	5 (2%)	28 (10%)	4 (1%)	12	299
≥2 years	86 (90%)	3 (3%)	4 (4%)	3 (3%)	1	97
Total	381 (85%)	11 (2%)	50 (11%)	8 (2%)	16	466

Note: Percentages exclude those with unknown NRTI backbone

Figure 8.8 - Initial NRTI backbone, by age at ART initiation and year of ART initiation, among those with data available from ART initiation
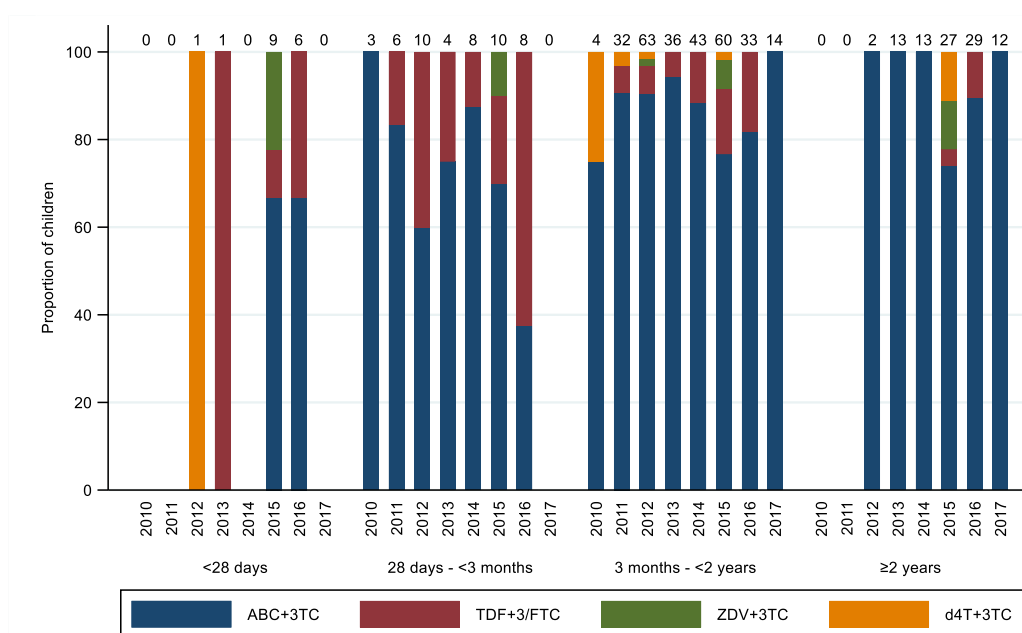
Notes: The number above each bar represents the number of children born in each year initiating ART in each age group. Includes only 447 with initial regimen of form 2NRTI + 1NNRTI/PI.

Initial regimens based on ARTemis data are shown in Table 8.13 and Table 8.14. A total of 216 children were included, the last of whom initiated treatment in May 2015. In terms of initial third agents, no children were reported to have initiated on NVP. The most common regimen for those <14 days was EFV-based (86%) and for those >28 days to 3 years was LPV-based (90%). For NRTI backbones, 93% (169/181) were reported to have been on ABC+3TC, with 10 on TDF+3/FTC (including 6/7 of those starting <28 days) and 2 on d4T+3TC. Comparing at an

individual-level, 127 children were linked between ARTemis and ACCDB, of whom 97 had a known initial regimen in both datasets, for whom the two regimens matched for 76 (78%).

Table 8.13 - Initial third agent, by age at ART initiation, based on data from ARTemis

Age at ART initiation	LPV/r	EFV	NVP	Unknown	Total
<14 days	1 (14%)	6 (86%)	0	1	8
14 - <28 days	0	0	0	0	0
28 days - <3 months	22 (92%)	2 (8%)	0	3	27
3 months - <3 years	133 (90%)	14 (10%)	0	32	179
≥3 years	1 (50%)	1 (50%)	0	0	2
Total	157 (87%)	23 (13%)	0	36	216

Table 8.14 - Initial NRTI backbone, by age at ART initiation, based on data from ARTemis

Age at ART initiation	ABC + 3TC	ZDV + 3TC	TDF + 3/FTC	d4T + 3TC	Unknown	Total
<28 days	1 (14%)	0	6 (86%)	0	1	8
28 days - <3 months	23 (96%)	0	1 (4%)	0	3	27
3 months - <2 years	136 (97%)	0	3 (2%)	1 (1%)	23	163
≥2 years	9 (90%)	0	0	1 (10%)	8	18
Total	169 (93%)	0	10 (6%)	2 (1%)	35	216

CD4 at ART initiation

Among the 466 children with data available from ART initiation, the proportion with an available CD4 measurement and their CDC stage at ART initiation is shown in Table 8.15, overall and by age at and year of ART initiation. 229 (49%) had a CD4 measurement available at the time of initiation, increasing with age from 44% to 55% to 75% among those aged <1 year, 1 to <6 years, and ≥6 years at ART initiation respectively. Over time, the proportion with a measurement decreased steadily from 70% among those starting treatment in 2010, to 38% in 2017. Overall, 67% (154/466) were CDC stage 3 (the most severe). Children starting ART at older ages had better immunological status (75% of those aged <1 year were CDC stage 3, decreasing to 63%, 56% and 33% among those 1 to <3, 3 to <6 and ≥6 years respectively), as did those starting in more recent years. Of 88 children who started ART aged older than one year between 2010 and 2013 and who had a CD4 measurement available, the median (IQR) CD4 count was 684 (414, 1129) (n=50) and CD4% was 17 (12, 23) (n=67); 17 (19%) did not meet the threshold for treatment start (≤750/≤25%).

Table 8.15 - CD4 tests conducted and CDC stage at ART initiation, overall and by age at and year of ART initiation, among those with data available from ART initiation

	N	Number (%) with a CD4 measurement	CDC immune stage		
			1	2	3
Overall	466	229 (49%)	35 (15%)	40 (17%)	154 (67%)
By age at initiation					
<1 year	245	107 (44%)	17 (16%)	10 (9%)	80 (75%)
1 - <3 years	168	92 (55%)	14 (15%)	20 (22%)	58 (63%)
3 - <6 years	49	27 (55%)	2 (7%)	10 (37%)	15 (56%)
≥6 years	4	3 (75%)	2 (67%)	0	1 (33%)
By year of initiation					
2010	10	7 (70%)	2 (29%)	1 (14%)	4 (57%)
2011	43	27 (63%)	2 (7%)	2 (7%)	23 (85%)
2012	84	52 (62%)	3 (6%)	4 (8%)	45 (87%)
2013	54	23 (43%)	4 (17%)	6 (26%)	13 (57%)
2014	66	34 (52%)	8 (24%)	9 (26%)	17 (50%)
2015	107	45 (42%)	8 (18%)	7 (16%)	30 (67%)
2016	76	31 (41%)	4 (13%)	8 (26%)	19 (61%)
2017	26	10 (38%)	4 (40%)	3 (30%)	3 (30%)

Proportions for the number with a CD4 measurement are of the total number in each group, and the proportions at each CDC stage are of those with an available CD4 measurement.

8.4.3. Outcomes after ART initiation

Among 619 children with follow-up after ART initiation, outcomes on ART are explored below.

<u>Follow-up status</u>

The follow-up status of all 619 children on ART in the sub-district is shown in Table 8.16, overall and separately for those who initiated ART in the sub-district and those who initiated ART elsewhere. The median (IQR) duration of follow up was 1.2 (0.3, 3.2) years. Two children had no ART follow-up after the date of initiation and were therefore excluded from the rest of the analyses in this section, leaving 617 children remaining.

Overall, at the time of data extract, 304 (49%) children were still in follow-up, 119 (19%) had transferred to another clinic, and 163 (26%) were lost-to-follow-up. The remaining 33 (5%) were reported to have died (of whom 31 died within 1 year of ART initiation), with an overall mortality rate of 2.7 (1.9, 3.8) per 100 person-years. Of the 33 children known to have died, 29 deaths were reported in ACCDB, and 4 were identified through linkage to demographic surveillance only. The follow-up status was broadly similar for those who had initiated ART within the sub-district compared to those who had initiated elsewhere, except that the proportion who had died or been lost-to-follow-up was slightly higher.

Table 8.16 - Follow-up status of all children on ART within the sub-district, by location of ART initiation

	Initiated ART in sub-district	Initiated ART elsewhere and later transferred in to sub-district	Total
	median (IQR) [range] or n (%)		
Total	466	153	619
Duration of follow-up (years)	1.2 (0.3, 3.2) [0.0, 6.8]	1.5 (0.5, 3.2) [0.1, 6.8]	1.2 (0.3, 3.2) [0.0, 6.8]
No follow-up after ART initiation	2	0	2
Follow-up status			
Still in follow-up	221 (47%)	83 (54%)	304 (49%)
Transferred out	79 (17%)	40 (26%)	119 (19%)
Lost-to-follow-up	134 (29%)	29 (19%)	163 (26%)
Death	32 (7%)	1 (1%)	33 (5%)

Among those with an available CD4 measurement at ART initiation, the proportion who were CDC stage 3 was 93% (14/15) among those who died and 72% (47/65) among those lost-to-follow-up, compared to only 61% (73/119) of those still in follow-up.

Of the total of 617 children, 558 (90%) were seen at only one clinic in the sub-district, 56 (9%) were seen at two, and 3 (<0.5%) were seen at three or more.

Frequency of viral load and CD4 monitoring

Among all 617 children on ART, 338 (55%) and 398 (65%) had any viral load and CD4 data recorded in ACCDB (including some measurements prior to ART initiation), and 277 (45%) and 246 (40%) had data after ART initiation respectively. Of the 311 (50%) with neither a CD4 or a viral load measurement after ART initiation, the median (IQR) [range] duration of follow-up on ART was 0.4 (0.1, 1.0) [0.0, 4.8] years compared to 3.0 (1.7, 4.5) [0.0, 6.8] among those with a measurement. Overall, including children with no measurements reported, the rate of viral load testing was 0.48 (95% CI 0.44, 0.52) per year (equivalent to approximately one test every two years), and of CD4 testing was 0.39 (0.35, 0.43) per year.

A summary of the rate of viral load and CD4 monitoring by clinic is shown in Table 8.17 and Figure 8.9. The rate of viral load testing per patient per year ranged from 0.15 to 0.87 across the clinics (p<0.001), and CD4 testing from 0.17 to 0.69 (p<0.001). There was evidence that smaller clinic size, measured using total follow-up time, was associated with the rate of CD4 testing (correlation=-0.616, p=0.009) but not viral load testing (correlation=-0.348, p=0.174). There was a strong positive association between the rate of viral load testing and the rate of CD4 testing at each clinic (correlation 0.70, p=0.002), as shown in Figure 8.10.

Table 8.17 - Rate of viral load and CD4 testing on ART, overall and by clinic

Clinic	Number of children ever seen	Total follow-up time (years)	Total number of viral loads conducted	Rate of viral load testing per year (95% CI)	Total number of CD4 tests conducted	Rate of CD4 testing per year (95% CI)
2	17	32.5	5	0.15 (0.06, 0.37)	11	0.34 (0.19, 0.61)
13	67	136.3	44	0.32 (0.24, 0.43)	27	0.20 (0.14, 0.29)
11	28	42.3	14	0.33 (0.20, 0.56)	16	0.38 (0.23, 0.62)
14	49	108.8	41	0.38 (0.28, 0.51)	19	0.17 (0.11, 0.27)
9	39	81.3	34	0.42 (0.30, 0.59)	21	0.26 (0.17, 0.40)
6	100	188.5	79	0.42 (0.34, 0.52)	62	0.33 (0.26, 0.42)
10	32	71.8	35	0.49 (0.35, 0.68)	38	0.53 (0.39, 0.73)
8	25	38.6	19	0.49 (0.31, 0.77)	20	0.52 (0.33, 0.80)
17	41	63.9	35	0.55 (0.39, 0.76)	39	0.61 (0.45, 0.83)
3	17	31.8	18	0.57 (0.36, 0.90)	22	0.69 (0.45, 1.05)
4	119	100.1	57	0.57 (0.44, 0.74)	37	0.37 (0.27, 0.51)
15	11	13.4	8	0.60 (0.30, 1.19)	7	0.52 (0.25, 1.09)
5	31	48.8	31	0.63 (0.45, 0.90)	26	0.53 (0.36, 0.78)
16	43	89.7	59	0.66 (0.51, 0.85)	43	0.48 (0.36, 0.65)
12	32	39.3	28	0.71 (0.49, 1.03)	17	0.43 (0.27, 0.70)
7	27	53.9	39	0.72 (0.53, 0.99)	34	0.63 (0.45, 0.88)
1	4	14.9	13	0.87 (0.51, 1.50)	10	0.67 (0.36, 1.25)
Overall	617	1,155.9	559	0.48 (0.44, 0.52)	449	0.39 (0.35, 0.43)
p				<0.001		<0.001

Note: p-value for difference in testing by clinics

Figure 8.9 - Rate of viral load (top) and CD4 (bottom) testing on ART, overall and by clinic
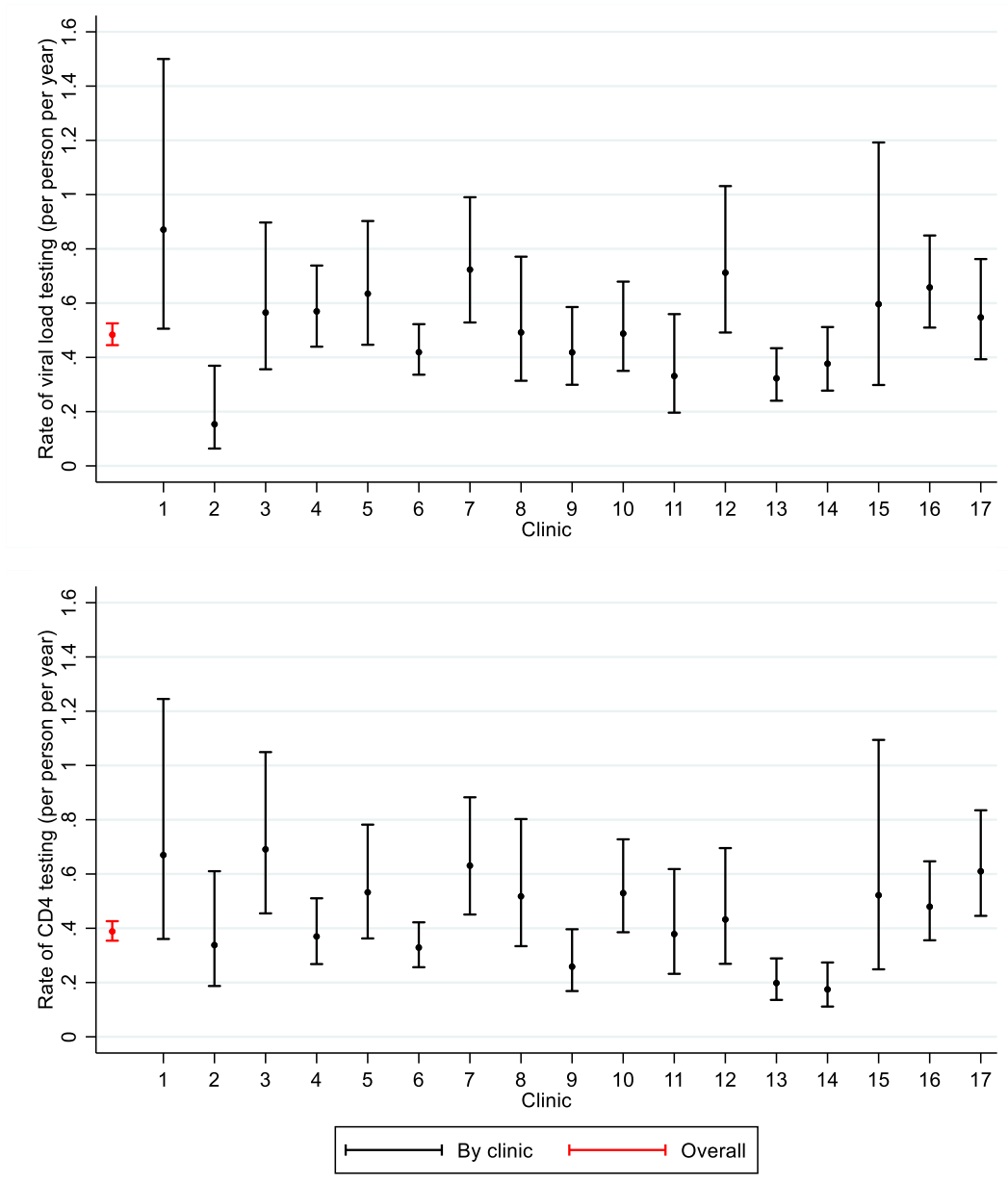
Figure 8.10 - Association between the rate of viral load testing and the rate of CD4 testing at each clinic
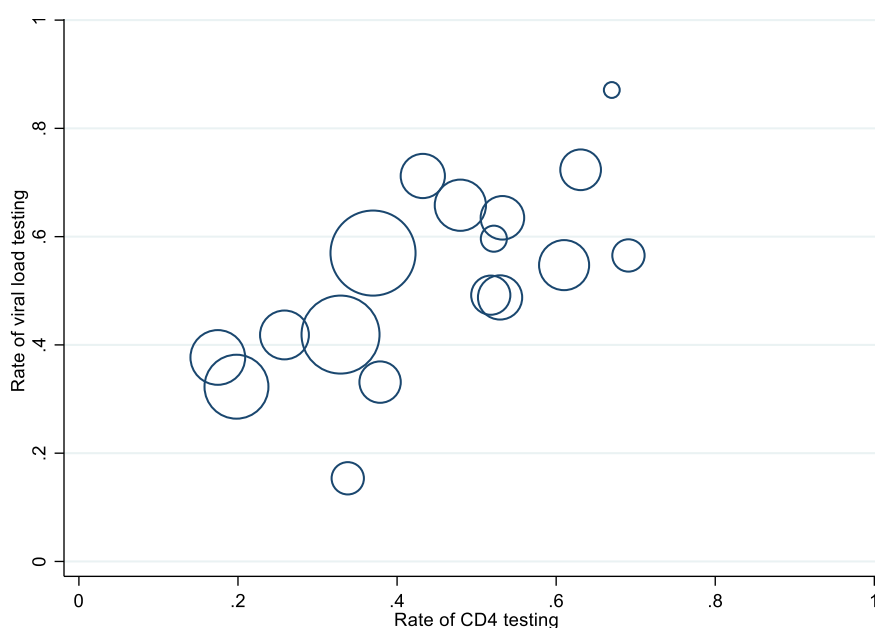
Note: Marker size represents the number of children seen at each clinic

The rate of testing of each marker is shown by time since ART initiation in Table 8.18 and Figure 8.11. There was no evidence of a change in the frequency of viral load testing by time since ART start (p=0.422), but CD4 testing decreased with increasing time on ART from 0.42 (0.35, 0.49) tests per year in the first year on treatment to 0.11 (0.03, 0.33) after five years (p<0.001).

Table 8.18 - Rate of viral load and CD4 testing on ART, overall and by time since ART initiation, among those with data from ART initiation

Years on ART	Number of children ever seen	Total follow-up time (years)	Total number of viral loads conducted	Rate of viral load testing per year (95% CI)	Total number of CD4 tests conducted	Rate of CD4 testing per year (95% CI)
<1	464	327.0	143	0.44 (0.37, 0.52)	136	0.42 (0.35, 0.49)
1-<2	258	210.9	121	0.57 (0.48, 0.69)	96	0.46 (0.37, 0.56)
2-<3	172	144.3	70	0.49 (0.38, 0.61)	60	0.42 (0.32, 0.54)
3-<4	120	101.7	52	0.51 (0.39, 0.67)	28	0.28 (0.19, 0.4)
4-<5	85	65.2	25	0.38 (0.26, 0.57)	14	0.21 (0.13, 0.36)
5-<6	46	28.1	7	0.25 (0.12, 0.52)	3	0.11 (0.03, 0.33)
6-<7	13	4.9	2	0.41 (0.10, 1.63)	0	-
Overall	464	882.0	420	0.48 (0.43, 0.52)	337	0.38 (0.34, 0.43)
p				0.422		<0.001

Note: p-value for change testing by time on ART

Figure 8.11 - Rate of viral load (top) and CD4 (bottom) testing on ART, overall and by time since ART initiation, among those with data from ART initiation
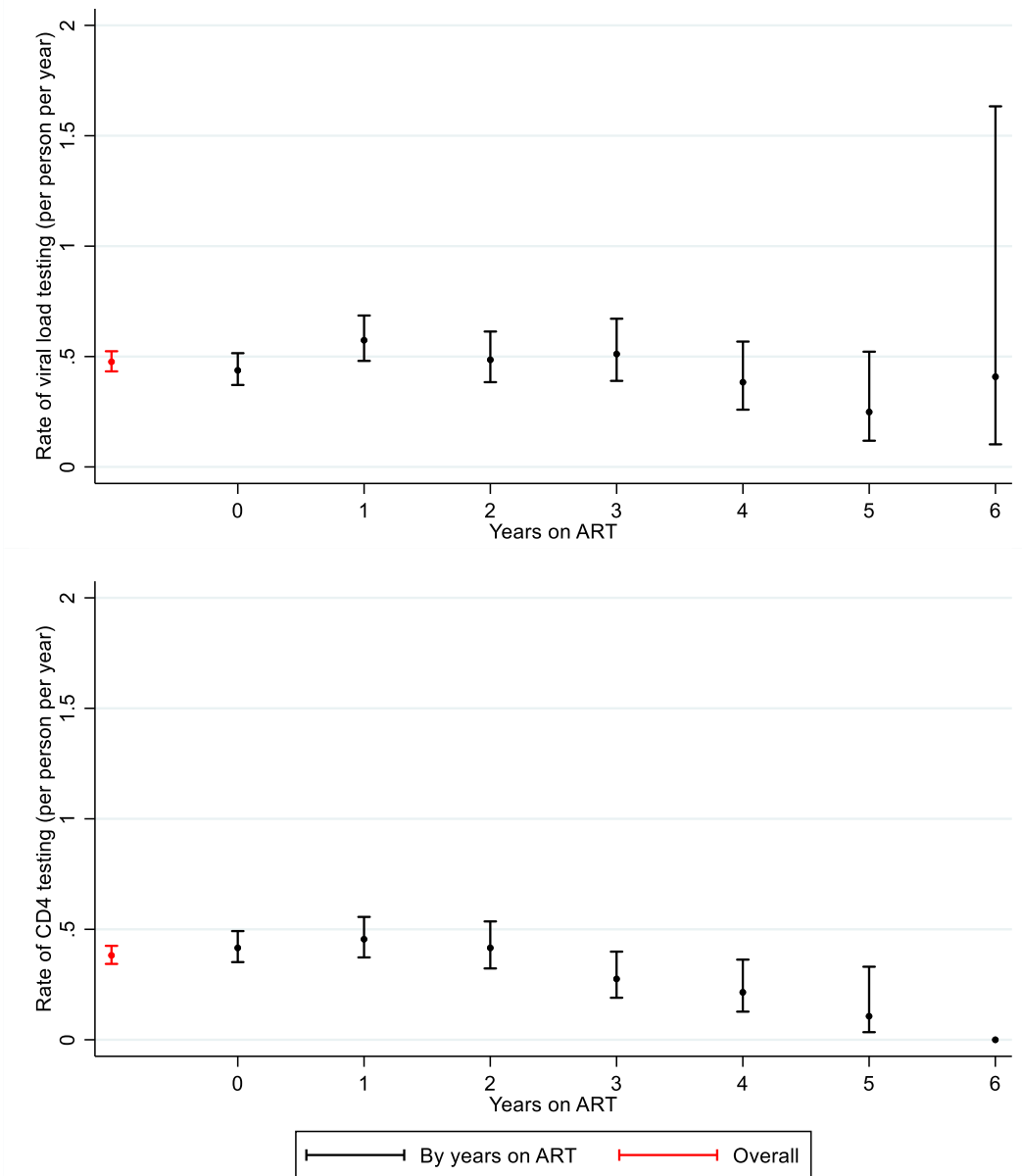
The rate of viral load and CD4 testing by calendar year is shown in Table 8.19 and Figure 8.12. The rate of viral load testing varied between 0.43 (0.35, 0.54) and 0.59 (0.50, 0.69) tests per year, with no evidence of trend over time (p=0.256), while the rate of CD4 testing decreased over time from 0.93 (0.23, 0.71) in 2010 to 0.15 (0.10, 0.22) in 2017 (p<0.001).

Table 8.19 - Rate of viral load testing by calendar year, overall and by calendar year

Calendar year	Number of children ever seen	Total follow-up time (years)	Total number of viral loads conducted	Rate of viral load testing per year (95% CI)	Total number of CD4 tests conducted	Rate of CD4 testing per year (95% CI)
2010	12	2.2	0	-	2	0.93 (0.23, 3.71)
2011	58	26.9	14	0.52 (0.31, 0.88)	14	0.52 (0.31, 0.88)
2012	149	88.0	42	0.48 (0.35, 0.65)	54	0.61 (0.47, 0.80)
2013	196	138.2	63	0.46 (0.36, 0.58)	70	0.51 (0.40, 0.64)
2014	250	188.6	82	0.43 (0.35, 0.54)	81	0.43 (0.35, 0.53)
2015	345	248.7	147	0.59 (0.50, 0.69)	100	0.40 (0.33, 0.49)
2016	394	310.1	167	0.54 (0.46, 0.63)	103	0.33 (0.27, 0.40)
2017	372	171.7	44	0.26 (0.19, 0.34)	25	0.15 (0.10, 0.22)
Overall	617	1,155.9	559	0.48 (0.44, 0.52)	449	0.38 (0.35, 0.42)
p				0.256		<0.001

Note: p-value for change in testing over calendar time

*Figure 8.12 - Rate of viral load (top) and CD4 (bottom) testing on ART, overall and by calendar year*

199
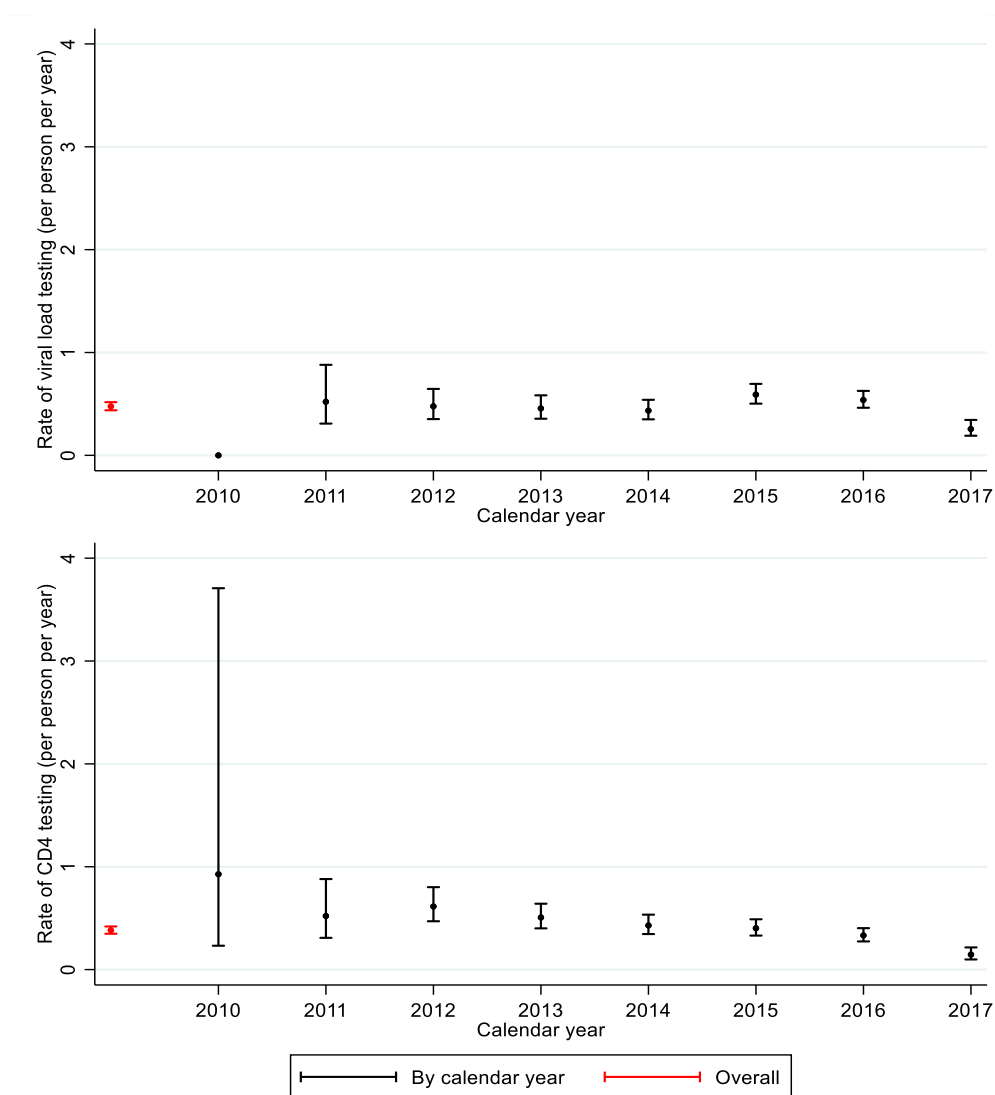
Of 617 children with follow-up on ART, 299 (48%) were linked to ARTemis. A comparison of the viral load and CD4 data in each dataset is shown in Table 8.20. Of the total number, 139 (46%) had viral load data recorded in either dataset, of whom 41 (30%) had a viral load measurement recorded in ACCDB that was missing from ARTemis and 56 (40%) had a viral load measurement recorded in ARTemis that was missing from ACCDB. Similar proportions were observed when comparing the availability of CD4 data.

Table 8.20 - Comparison of viral load and CD4 measurement data recorded in ACCDB and ARTemis

	Viral load	CD4
Total number of children linked between ACCDB and ARTemis	299	299
Number (%) with no measurement in either dataset	160 (54%)	136 (45%)
Number (%) with a measurement in either dataset	139 (46%)	163 (55%)
Measurement in ACCDB that was missing from ARTemis	41 (30%)	50 (31%)
Measurement in ARTemis that was missing from ACCDB	56 (40%)	64 (39%)

Viral suppression

Among the 464 children with data from ART initiation, 242 (52%) had any viral load measurements and 200 (43%) had viral load measurements after ART initiation. Among those with measurements after initiation, the first viral load was at a median of 9.0 (5.8, 16.7) months after ART initiation, at which time 58% (115/200) were suppressed. At 12 months after ART initiation, 15% (68/464) of children had a viral load measurement, of whom 66% (45/68) were suppressed.

Overall, including children who transferred in to the sub-district on ART, 77% (212/277) ever had a viral load ≤400 copies/mL (including only those with viral measurements reported). At their most recent viral load, 70% (194/277) of children were suppressed; in sensitivity analysis when assuming the 340 those with no measurements were suppressed this rose to 87% (534/617), and when assuming they were all unsuppressed fell to 31% (194/617). Those currently not on ART were less likely to be suppressed compared to those on ART (20/39 (51%) vs. 174/238 (73%), p=0.006), but among those on ART, there was no difference in the proportion suppressed by current regimen (p=0.553). There was no evidence of an association between suppression and current age (p=0.801).

Time to viral suppression after ART start is shown in Table 8.21. By 1 year on ART 26.6% had achieved viral suppression rising to 48.7% by 2 years, with higher proportions observed when restricting analysis just to those with any viral load data reported.

Table 8.21 - Time to viral suppression ≤400 copies/mL following initiation of ART

	Including all children (N=464)	Restricted to those with any viral load measurements reported after ART initiation (N=200)
	Proportion who achieved viral suppression (95% confidence interval)	
6 months	8.1% (5.6%, 11.5%)	13.8% (9.7%, 19.4%)
1 year	26.6% (21.2%, 32.0%)	41.4% (34.8%, 48.7%)
2 years	48.7% (42.3%, 55.4%)	66.3% (59.5%, 73.1%)
3 years	60.8% (53.6%, 68.0%)	76.8% (70.2%, 82.9%)

Of 110 children who ever had a viral load >1,000 copies/mL recorded, 69 (63%) ever had a repeat viral load test, which occurred at a median 33.7 (20.1, 62.0) weeks later, at which 33 (48%) were suppressed ≤400 copies/mL.

Immunological outcomes

Including both those with data from initiation and transferring in, 246 (40%) had any CD4 measurements on ART. At most recent CD4 measurement, 77 (31%) children were CDC stage 1, 86 (35%) were stage 2 and 83 (34%) were stage 3 (the most severe). Of the subset of 117 children with both a measurement at ART initiation and a subsequent measurement on ART, 18 (15%) had worsened to a more severe CDC stage, 50 (43%) remained at the same stage, and 49 (42%) had improved stage.

Interruptions to ART

Among the 617 children on ART, 392 (64%) children had one or more interruptions to treatment of ≥1 day, with the median (IQR) duration of each gap in treatment 33 (25, 50) days. Overall, 387 (63%), 319 (52%) and 39 (6%) of children had an interruption of ≥1 week, ≥1 month, and of ≥6 months respectively, and 75 (12%), 39 (6%) and 77 (12%) children had 2, 3 and ≥4 interruptions lasting ≥1 month respectively.

The medication possession ratio is shown in Figure 8.13. The median (IQR) ratio was 92.5% (76.6%, 100.0%), and 142 (23%) and 54 (9%) children were off treatment at least a quarter and half the time respectively (corresponding to a medication possession ratio of ≤75% and ≤50% respectively). By clinic, the median medication possession ratio ranged from 70.0% to 100.0%, with evidence of variation by clinic (p<0.001) (Table 8.22). There was no evidence of an association between the size of the clinic (measured using the number of children ever seen there) and medication possession ratio (p=0.531). The median prescription duration was 30 (28, 32) days.

Figure 8.13 - Medication possession ratio
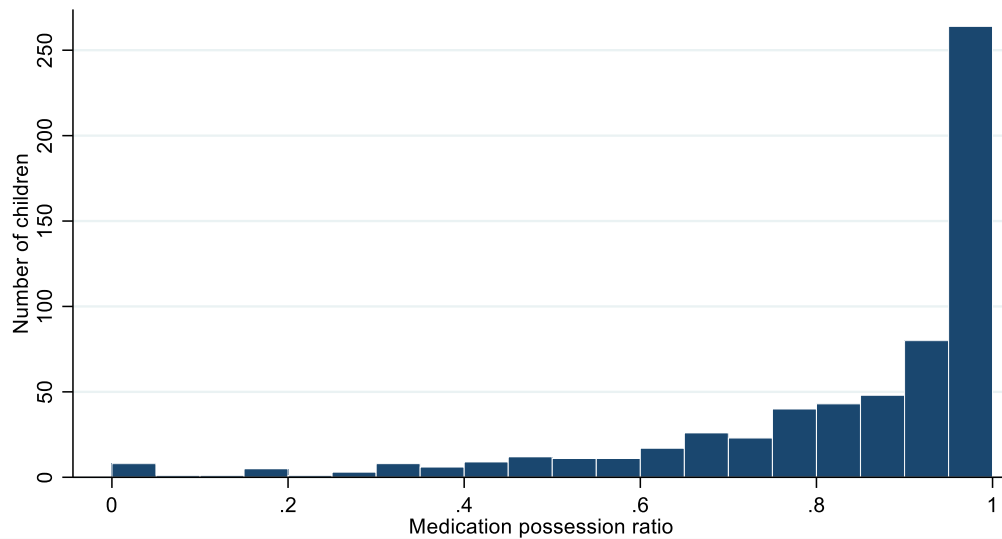
Note: Includes both children with and without (represented in the 95-100% bar) any treatment interruptions

Table 8.22 - Medication possession ratio, overall and by clinic

Clinic	Number of children seen at clinic	Medication possession ratio, median (IQR)
4	117	100.0% (100.0%, 100.0%)
12	32	100.0% (84.8%, 100.0%)
17	41	100.0% (94.2%, 100.0%)
13	67	96.5% (87.4%, 100.0%)
15	11	96.3% (76.7%, 100.0%)
6	99	93.2% (75.7%, 100.0%)
16	43	92.8% (85.3%, 100.0%)
7	26	91.7% (84.6%, 97.9%)
9	39	89.4% (63.6%, 100.0%)
5	31	86.5% (62.0%, 100.0%)
10	32	85.7% (79.0%, 100.0%)
2	17	84.9% (70.2%, 94.3%)
8	25	83.3% (70.7%, 100.0%)
14	49	82.2% (69.0%, 100.0%)
3	17	81.8% (66.1%, 92.2%)
11	27	81.5% (58.7%, 100.0%)
1	4	70.0% (61.2%, 83.5%)
Overall	617	92.5% (76.6%, 100.0%)
p		<0.001

Note: p-value for difference in testing by clinics

Treatment data for a random sample of children with a treatment interruption are shown in Figure 8.14 to give an idea of the frequency in which individual children were experiencing interruptions and the length of each interruption, with time on ART shown in black and off ART shown in pink. Although some children had extended periods without treatment, the majority had short and often frequent breaks.

Figure 8.14 - Time on and off ART for a random sample of children who interrupted ART treatment at least once and had data available from ART initiation

Note: Plot shows ART data for a random sample of 15% of children with any interruption to treatment and who had data available from ART initiation, with each horizontal line representing the ART data for one child

<u>Treatment changes</u>

Among those with data from ART initiation and known initial regimen (n=450), 137 (30%), 94 (21%) and 95 (21%) had any change to their regimen, any change to their third agent and any change to their NRTI, respectively (Table 8.23). When treatment changes only recorded at one visit were excluded (in instances where the child was reported to have switched back to their previous regimen at the next visit), the total number of treatment changes observed only decreased slightly, from 137 (30%) to 126 (28%).

Table 8.23 - Children who changed treatment, by type of agent changed and number of changes

	Any change	Third agent change	NRTI change
Number with a change	137 (30%)	94 (21%)	95 (21%)
Number of changes			
1	53 (12%)	56 (12%)	31 (7%)
2	42 (9%)	20 (4%)	42 (9%)
3	25 (6%)	10 (2%)	12 (3%)
≥4	17 (4%)	8 (2%)	7 (2%)
Total number of change events	288	162	194
Sensitivity analysis: ignoring changes recorded at only one visit:			
Number of children with a change	126 (28%)	88 (20%)	84 (19%)
Number of changes			
1	60 (13%)	59 (13%)	34 (8%)
2	38 (8%)	18 (4%)	34 (8%)
3	19 (4%)	7 (2%)	9 (11%)
≥4	9 (2%)	4 (1%)	7 (2%)
Total number of change events	235	135	160

41 (9%) children switched to second-line treatment. Of those initiating on an LPV/r-based regimen, 30/308 (9%) switched to second-line (all to an EFV-based regimen), among whom the median (IQR) age at switch was 3.7 (2.4, 5.0) years. Of those initiating on an EFV-based regimen 11/136 (8%) switched to second-line (all to an LPV/r-based regimen), at a median age of 3.3 (1.6, 4.6) years. Of 3 children initiating on a NVP-based regimen, none were reported to have switched.

Time to switch to second-line is shown in Figure 8.15. The median time from ART initiation to first switch was 1.9 (0.7, 3.4) years. The cumulative incidence of switch to second line by 1, 3 and 5 years after initiation was 3.9% (2.3%, 6.8%), 12.4% (8.5%, 17.9%) and 25.8% (18.5%, 35.3%) respectively. In unadjusted analysis, the risk of switch to second-line increased with age at ART initiation (HR 1.02 (95% CI 1.01, 1.05), p=0.045 per month older) and with year of initiation (1.42 (1.10, 1.82) per year later, p=0.007). There was no evidence of a difference in the risk of switch

among those who initiated on an EFV-based regimen compared to a LPV/r-based regimen (1.25 (0.62, 2.52), p=0.525).

Figure 8.15 - Time to switch to second-line
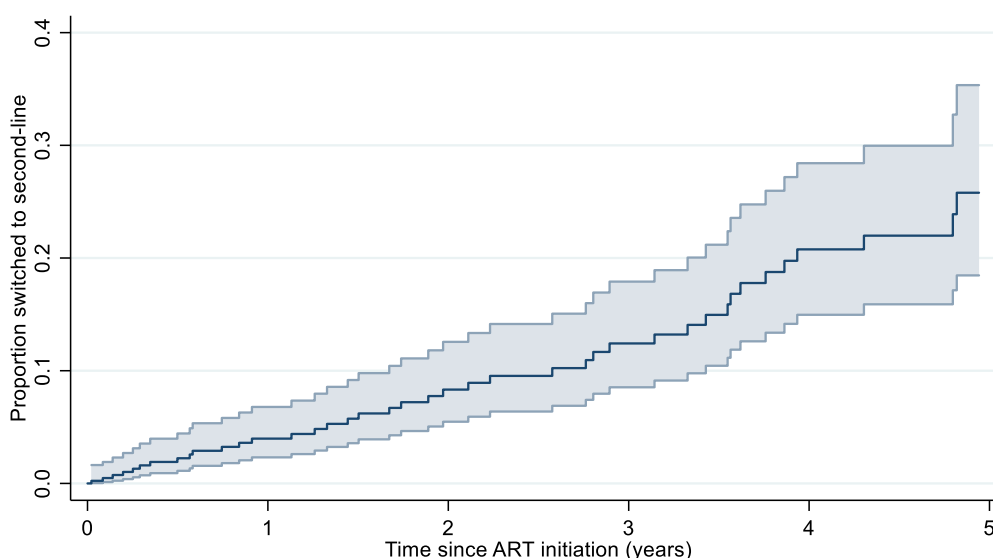
40 (59%) switches to second-line occurred after a break from treatment, with a median 59 (40, 88) days off ART before the switch, and 3 (4%) switches coincided with change of clinic. Of 41 children who switched to second-line treatment, 6 (15%) had a viral load available at the time of switch, of whom 4 (67%) were suppressed, and 5 (7%) had a CD4 measurement available, of whom 1 (20%) was CDC stage 1, 2 (40%) were stage 2, and 2 were stage 3 (40%).

8.5. Discussion

Across the whole time period, the estimated ART coverage among children with a positive PCR test was 53%, with those born in more recent calendar years more likely to initiate treatment. Correspondingly, the median time from diagnosis to ART initiation fluctuated between 7.1 and 8.9 weeks between 2010 and 2013, after which it fell steadily over time to only one week among those born in 2016. The proportion initiating treatment was lower among those born in 2016 than 2015 which may have been the result of the reduced follow-up time available for this group; the interpretation of an analysis restricted to initiations within 6 months of diagnosis was limited by small numbers of those born in 2016. Following guidance from WHO, national treatment guidelines changed in 2013, at which point immediate ART to all children under 5 years of age was recommended, regardless of CD4 count. Although the majority of those included in this analysis were under 1 year of age at time of HIV diagnosis and thus would have been eligible for immediate ART before and after this guideline change, the increase in coverage after 2013 suggests that the previous guidelines may not have been well understood or implemented by healthcare workers. The proportion of children with severe immunosuppression at initiation decreased over time, likely a reflection of the quicker time to ART initiation, however was still

over 60% in 2016. This late presentation to care has been shown to be associated with poorer outcomes long-term [254].

The majority of studies assessing ART coverage among children with HIV in South Africa are based at single hospitals, often in an urban area, where coverage is likely to be higher due to better resources within healthcare facilities [255], with estimates ranging from 71% to 98% [165, 182, 184]. Few studies have considered linkage to ART in rural areas such as this setting. One study which assessed ART coverage among children diagnosed at a hospital in another part of uMkhanyakude district in 2012 reported an estimate of 56%, closer to that shown here [175].

There are several possible explanations for the low estimated coverage of ART. Firstly, it may be a consequence of the relatively high levels of migration and mobility in the sub-district; when the analysis was restricted to the subset of children known to be still resident in the area a year after their diagnosis, the proportion linking to ART increased to 70%. Secondly, among those who did not start ART and were linked to ACDIS, 43% died. These children represent failures of the current healthcare system, and demonstrate the high mortality among children with HIV who do not receive treatment, which has been well documented [62]. Finally, it is likely that some children who did in fact start ART were not linked between NHLS and ACCDB. A large number of children born in the time period of interest were recorded as being on ART within the sub-district but were not linked to a positive PCR test result. Many of these children were likely either to have been diagnosed outside of the sub-district or to have been diagnosed at an older age with an antibody test. However, the remaining 30% may actually overlap with the group who had a positive PCR test result but were not found to be on ART, as the PCR test and ART data for some children may not have been successfully linked. The upper limit of the potential impact of this is an increase in the estimate of ART coverage from 53% to 77%, which would be more in line with the results from other studies.

Only 46% of children diagnosed with HIV at birth initiated ART, compared to 73% of those diagnosed at older ages. The lower ART coverage might be explained by the fact that linking PCR test records conducted at birth may be more difficult than linking those conducted at older ages, as the use of mother's name in place of child's would be more likely at birth. The difference may also be the result of survivor bias, given the early peak of mortality in this group. Only two other studies have considered the impact of the introduction of birth testing in South Africa on the subsequent uptake of ART, both at large urban hospitals. One found 86% of those diagnosed at birth were linked to HIV care compared to 100% of those diagnosed later [165]. In the other, 95/99 (96%) of children diagnosed at birth initiated ART (with no comparison group), however active outreach was used to trace these children, and only 3,251/6,261 (52%) of those who tested negative at birth returned for their results [163]. In my analysis, those diagnosed at birth who did go on to initiate treatment did so quicker and at a much earlier age however (median

1.4 weeks vs. 19.7 among those diagnosed later), with recent evidence suggesting this very early initiation is associated with a reduction in the size of the viral reservoir [256].

Among those starting ART, results suggested the use of initial regimens not usually recommended for young children. NVP-based regimens were not commonly used in neonates aged <14 days, with all but one receiving either a LPV/r or EFV-based regimen, despite these drugs not being licensed for this age group due to toxicity concerns [68]. Around a quarter of younger children aged up to 3 months used TDF, even though this drug is not recommended until at least 2 years of age [68]. Additionally, a small number of children were initiated on a d4T-containing regimen until 2015, despite South African guidelines from 2010 stating that no children should newly start this drug, and from 2013 that all children already taking it should be switched to an alternative drug. It is likely that some of these unexpected findings reflect poor data quality in TIER.net (and thus ACCDB), rather than true initial regimen choices. For example, no generic paediatric formulations of TDF are available [257], making its use here implausible. Another study in South Africa found that 383 children (aged <15 years) were reported to be on TDF, and chose to exclude them because of presumed misclassification [258]. Initial regimens reported in ARTemis, which is based on data collected by AHRI staff specifically for research, were more in line with recommendations especially for children older than 1 month, although still none of the children initiating <14 days were reported to be on an NVP-based regimen, as has been reported in other studies from South Africa [184].

The rate of viral load monitoring was 0.48 tests per year on ART, equivalent to approximately one every 2 years. This is much lower than the rates recommended in the South African guidelines, which were (for children aged under 5) at least six-monthly viral loads until 2015, and then annually thereafter. The frequency of testing varied by clinic from between 0.15 and 0.87 tests per child per year, but it is unclear whether this reflects variation in actual testing practices or variation in completeness of data entry into TIER.net. In a comparison of recording of viral load data, 40% of children had a measurement recorded in ARTemis that was missing in ACCDB, suggesting poor data entry. Of note, a similar proportion had a measurement missing in the other direction (that is, in ACCDB but not ARTemis), suggesting ARTemis was also not complete. A study across four districts in South Africa found that the proportion of patients with a viral load test ranged between 71% and 90% by district, with only 86% of these viral load tests recorded in TIER.net [259]. Although routine monitoring of all children is recommended, it is possible that testing was more targeted, with only those failing clinically and thus suspected of poor adherence or virological failure tested, which would present a selection bias. The frequency of CD4 testing was similarly low, and decreased over calendar time, in line with the move away from routine CD4 monitoring among stable, virologically suppressed patients [260].

The potential selection bias in viral load testing and infrequency of viral load measurements make it hard to interpret the proportion of children achieving viral suppression. Analysis of the most recent viral load for each child suggested that 70% of tested children were suppressed. This may be an underestimate of the true proportion if viral load tests were used in a targeted way, with an upper bound (assuming all those not tested were suppressed) of 87%. The estimate is however in line with other studies of children in South Africa, which range from 56% to 88% [165, 184, 261], with children often taking longer to suppress than older children and adults [262].

Over half of children experienced an interruption to treatment of over one month. Reasons for stopping and starting ART are not available in ACCDB, but the fact that the majority were short interruptions suggests that they could be the result of delays in attending clinic for ART prescription refills, rather than being planned interruptions. The median duration of ART prescriptions among this group was only one month, and regular visits to clinic such as this represent a huge burden for caregivers, especially given the long median travel time and cost of public transport [135]. It is also possible that interruptions were the result of stock-outs; a cross-sectional study in South Africa in 2015 showed 9% of clinics reported a stock out of at least one paediatric ART drug in the last 3 months [263]. It is also possible that these short interruptions were a result of data entry error. Unplanned treatment interruptions are not generally recommended in young children because of concerns over rapid disease progression off ART, especially in settings with limited access to regular viral load monitoring [264]. Comparative data on treatment interruptions among children of the same age in South Africa are not available, however a study of children and adolescents aged up to 20 years (but not disaggregated by age) which used data from clinic notes reported fewer interruptions than observed here; 8% had an interruption of at least 7 days in the first 24 months on ART in 2009, decreasing to 5% in 2013 [265]. It is also possible that during periods of treatment interruption, the children could have been receiving ART at another clinic, either within or outside of the sub-district, but the records from the different clinics had not been linked within ACCDB by AHRI. The data presented here only represent time that children were without an ART prescription, and do not include additional time not taking ART; no data on actual adherence were available, and the infrequency of viral load measurements makes this difficult to infer.

In this analysis, the cumulative incidence of switch to second-line by 3 years after ART initiation was 12.4%. Many of these switches were from LPV/r to EFV at around 3 years of age, which may have been for simplification of administration rather than a true switch to second line. These switches may also have been made to avoid a drug-drug interaction between LPV/r and rifampicin in those initiating treatment for tuberculosis [266], or be the result of misinterpretation of guidelines, which recommend EFV for children newly initiating after age 3

years rather than for all children over 3 years. There are no studies of switching in children specifically of this age group and in South Africa, but this is higher than estimates from other studies across sub-Saharan Africa and in older children. A study of children and adolescents aged 3 to 20 years at clinics in Johannesburg reported that 3.3% switched to second-line by 2 years on ART [265]. Among 8,082 children from South Africa and Botswana who initiated ART before the age of 3 years, the cumulative incidence of switch by 3 years after ART initiation was 3.7% [267]. It is possible that data entry errors led to overestimation of the number of treatment changes, especially since the TIER.net system requires regimens to be re-entered at each clinic visit, although when regimens recorded at only one visit were excluded there was little change to the results.

By last follow-up, 5% of children had died and 26% had been lost-to-follow-up. Mortality here was lower than another study of children in South Africa, which reported that 12% had died and 15% were lost-to-follow-up by 1 year after initiation. The completeness of recording of death data in TIER.net is unclear, with some deaths among those on ART not recorded here but identified only through demographic surveillance. It is possible that mortality is actually higher among these children classified as lost-to-follow-up, as observed in other studies [268], with severe immunosuppression more prevalent among those who were lost-to-follow-up compared to those who remained in care.

8.6. Key findings

The key findings from this chapter are:

- ART coverage among children diagnosed with HIV was 53%. Those born more recently were more likely to start, however a high proportion still initiated with severe immunosuppression, demonstrating late presentation to care.
- A lower proportion initiated among those diagnosed at birth compared to later, although this may be a result of difficulties in linking birth PCR tests to other data.
- Among children who were diagnosed but did not start ART (and were linked to ACDIS), 43% were known to have died
- The use of initial regimens not normally recommended in younger children was reported, which may reflect poor data quality of TIER.net and ACCDB.
- Half of children experienced an extended interruption to ART of over one month.
- A low proportion of children had any viral load data recorded, but among those who did, 70% were suppressed at their most recent test.
- By 3 years on ART, the cumulative incidence of switch to second line was 12.4%.
- By last follow-up, 26% of children were reported lost-to-follow up and 5% had died.

Chapter 9. The cascade of care

9.1. Introduction

In this chapter, I bring together results from the previous chapters to estimate the cascade of care in the sub-district. As with Chapter 8, because of the longer duration of follow-up included in this chapter, the term 'children' is used throughout rather than 'infants', however all analyses are still based on the same group of individuals born in the sub-district between 1st June 2010 and 31st December 2016.

I estimate the cascade in two different ways; firstly, the cascade stage of all children with HIV still living in the sub-district on the 31st June 2017, and secondly, the cascade stage at 2 years of age of all children exposed to HIV. In order to estimate these cascades, I additionally use estimates of the rate of MTCT in South Africa to estimate the total number of children born in the sub-district who acquired HIV.

9.2. Objectives

The objectives for this chapter are to:

- Estimate the number of children born in the sub-district who acquired HIV, and the proportion of these who were never diagnosed
- Estimate the cascade of care for children born in the sub-district, firstly restricted to those who acquired HIV, and secondly for all HIV-exposed children

9.3. Methods

9.3.1. Estimating the number of children born in the sub-district who acquired HIV

In order to estimate the cascade of care, an estimate of the total number of children born within the sub-district who acquired HIV (either in utero, intrapartum or postpartum) was required. This was calculated as the number of HIV-exposed children born in the sub-district multiplied by an estimate of the rate of MTCT, and was calculated both overall and by year of birth.

The number of HIV-exposed children born in the sub-district was taken from Chapter 5. Of the four methods used to estimate PCR testing coverage, only methods 2 and 3 generated estimates of the total number of HIV-exposed children born in the sub-district, with the others looking at coverage in a smaller region within the sub-district. With only method 2 (which was based on the number of live births as reported by SSA and the antenatal seroprevalence as estimated by

ANCHSS) was it possible to estimate the number born over the whole time period of interest (June 2010 to December 2016), and this estimate was therefore used in analysis here.

Estimates of the rate of MTCT over time were taken from Thembisa, a mathematical model of the South African HIV epidemic [269]. Estimates of the rate of MTCT to 18 months of age, that is, including those who acquired HIV through breastfeeding, are produced for each calendar year at both a national and provincial level, and so the estimates specific to KwaZulu-Natal were used here.

Two other sets of estimates of the rate of MTCT in South Africa were available, and for comparison, the number of children with HIV was additionally calculated based on these. The first set of estimates came from Spectrum and the AIDS Impact Module (AIM), a system of mathematical models that support analysis, planning and advocacy for health programs [270], and which includes estimates of the rate of MTCT. Estimates came from Spectrum version 5.756, and were based on the default input parameters provided for South Africa. The second set of estimates were taken from UNAIDS [39]. These were also based on Spectrum, but the model was updated following recommendations from the UNAIDS Reference Group on Estimates, Modelling and Projections, with the main change made being a reduction to the transmission probabilities for women seroconverting during pregnancy. Estimates were not available for 2010, 2012 or 2016; for 2010 and 2012 the average of the estimates from the years before and after was used, and for 2016 the rate was assumed the same as estimated for 2015. The Thembisa estimates were used for the main analysis, as they were based on a model designed specifically for South Africa and KwaZulu-Natal.

Two other sets of estimates of the rate of MTCT are also presented for comparison, although they were not used to calculate estimates of the total number of children with HIV. Firstly, SAPMTCTE (South African Prevention of Mother to Child Transmission Evaluation) was a survey of 2,877 HIV-exposed infants born between October 2012 and May 2013, which enrolled from sites across South Africa (including KwaZulu-Natal) [271, 272]. Infants received PCR tests as recommended in the national guidelines and then at 3 month intervals up to 18 months of age. The resulting incidence rate was weighted by the authors to account for loss-to-follow-up and for eligible infants for whom consent was not given. For comparison with the estimates from the other methods which were presented by calendar year of birth, the SAMPTCTE estimate was summarised as 2013. Given the short time period covered, this was not used to estimate the number of children with HIV in my analysis. Secondly, the proportion of infants with a positive PCR test in my analysis (see Chapter 7) was summarised. This was likely to be an underestimate of the true transmission rate given the low testing coverage and low proportion of infants with a repeat test, and so was not used in analysis here.

From the estimates of the number of children with HIV calculated based on Thembisa, Spectrum and UNAIDS, the number of undiagnosed infections was calculated as the total number of children with HIV minus the number diagnosed. The number of children born in the sub-district who were diagnosed was calculated as the number who either had a positive PCR test and/or who were on ART and assumed born in sub-district because either they initiated ART at one of the clinics in the sub-district or because they had received a negative/indeterminate PCR test at one of the clinics in the sub-district. That is, those who transferred into a sub-district clinic on ART and who had no linked PCR tests were assumed to have been born elsewhere and therefore excluded.

9.3.2. The cascade of care

The cascade of care can be constructed in many different ways, including variations on the population of interest, the time point (or time points) at which measurement will be made, the specific stages to be included, and the sub-groups by which the cascade will be summarised. For this analysis, two different cascades were estimated. The methods for each are described below and shown in Table 9.1, followed by a description of the definitions used for the stages across both cascades.

The first cascade (cascade #1) was constructed cross-sectionally and designed in line with the UNAIDS 90-90-90 targets, as a snapshot of the status of individuals on the 30th June 2017. The denominator was all children born in the sub-district between 1st June 2010 and 31st December 2016 who had acquired HIV, who were not known to have died (through either ACDIS or ACCDB), and who were not known to have transferred ART care to a clinic outside of the sub-district through ACCDB (that is, were still assumed to be living in the sub-district). Although data on migration were available for the subset of children who were linked to ACDIS (i.e. including some of those not on ART), this could not be used as this would have required the time-consuming classification of free text descriptions of their destination. The follow-up status of each child as of the 30th June 2017 (chosen as both PCR test data from NHLS and ART data from ACCDB were available to mid-July 2017, with the exact date varying by clinic) was assessed. The inclusion criteria correspond to including all children with HIV who were born and still living in the sub-district, and who were aged under 7 years and 1 month on 30th June 2017. Presenting results of this cascade over time, for example summarising the follow-up status of those living with HIV in the sub-district at the end of each calendar year, was considered; however the increasing follow-up time would have made the interpretation of changes over time difficult, and so this approach was not taken. The cascade stages included were the proportion diagnosed with HIV, the proportion currently on ART, and the proportion currently virally suppressed.

The second cascade (cascade #2) aimed to determine the success of the sub-district's early infant diagnosis care program, up to 2 years of age. All HIV-exposed children born in the sub-district between the 1st June 2010 and 30th June 2015 were included, and the status of each was summarised at 2 years of age. Those born after this date were excluded, as they did not have enough follow-up time available. The proportion of children who received a PCR test (in the case of children who did not have HIV) or who were diagnosed with HIV by 2 years of age was summarised, followed by, for those diagnosed with HIV only, the proportion who initiated ART, their current ART status, and their current viral suppression status. Results were presented overall and stratified by year of birth.

Table 9.1 - Methods used for the analysis of each cascade

	Cascade #1	Cascade #2
	Current status of children alive and living with HIV in the sub-district as of 30th June 2017	Status at 2 years of age of all children born in the sub-district and exposed to HIV
Denominator	All those born in the sub-district between 1st June 2010 and 31st December 2016 who had acquired HIV, and who were still living in the sub-district	All those born in the sub-district between 1st June 2010 and 30th June 2015 who were exposed to HIV
Excluded	Those known to have died (through either ACDIS or ACCDB) or known to have transferred ART care to a clinic outside of the sub-district (through ACCDB)	None
Time point cascade analysed at	30th June 2017	2 years of age, summarised overall and by year of birth
Cascade stages included	• Diagnosis (classified at diagnosed/not diagnosed) • Current ART status (classified as: on ART/off ART) • Current viral suppression status (classified as: suppressed/not suppressed/no viral load available)	• Diagnosis (classified at diagnosed/not diagnosed) • Initiation of ART (classified as: initiated/died before initiation/did not initiate) • Current ART status (classified as: on ART/off ART/transferred out/died after ART start/lost-to-follow-up/unknown) • Current viral suppression status (classified as: suppressed/not suppressed/no viral load available)

The following definitions were used for each stage (where applicable) across both cascades:

- Number of HIV-exposed children: This was assumed to be as estimated using method 2 in Chapter 5, based on the number of live births as reported by SSA and the antenatal seroprevalence as estimated by ANCHSS.

- Number of children who acquired HIV: This was based on estimates of the rate of MTCT from Thembisa multiplied by the number of HIV-exposed children.

- Diagnosis: This was defined as the number with either a positive PCR test conducted at a clinic in the sub-district, or who were on ART and had either initiated treatment or had a PCR test conducted in the sub-district (to suggest that they hadn't been born elsewhere). The date of diagnosis was defined as the earliest of the date seen and the date of their first positive PCR test. The HIV status of children linked to ACDIS and known to have died without having received a PCR test was imputed, based on the estimates of the rate of MTCT.

- Initiation of ART: This was based on those with a date of initiation recorded, with this defined as the date first reported to be on ART in ACCDB. For 60 children known to be on ART but with no ART data available (i.e. with no dates on ART available, as described in Chapter 8), the date of initiation (and date of diagnosis if not otherwise known) was imputed based on the median age at initiation among children with the same year of birth. For the subset of children linked to ACDIS, data on death prior to ART initiation were available, with these individuals classified as having died prior to ART initiation in cascade #2 and excluded from the denominator of cascade #1.

- Current ART status: For cascade #2 this was categorised as either on or off ART (according to the visit data reported in ACCDB), lost-to-follow-up, died, or transferred out to another clinic (as reported in ACCDB), or unknown (for those with no follow-up data available). For cascade #1, this was categorised as either currently on or off ART.

- Viral suppression: Among those on ART, viral suppression was defined as having a viral load ≤400 copies/mL, allowing a window of ±6 months around each time point of interest (as long as the test was conducted after ART initiation). It was assumed that viral load data were missing at random, that is, that the same proportion of children were suppressed among those without a viral load as observed among those with a viral load.

An analysis of each cascade restricted to children who were linked to ACDIS was considered in order to be able to fully account for death and migration across whole cascade (rather than just after ART initiation), but was not conducted due to the small numbers of children eligible (for example, only 173 children with HIV).

9.4. Results

9.4.1. Estimating the number of children born in the sub-district who acquired HIV

The five sets of estimates of the rate of MTCT are presented in Table 9.2 and Figure 9.1. Based on the Thembisa model, the rate was estimated to be 11.06% in 2010 and decreased steadily over time to 4.37% in 2016. Estimates from Spectrum were the highest of the five for all years

of birth, ranging from 12.32% to 9.85%, with no clear trend over time. The UNAIDS estimates decreased from 9.00% in 2010 to 2.00% in 2016, and were approximately the same as the proportion with a positive PCR test, especially for later years. The estimate from SAPMTCTE for 2013 was slightly higher than those from UNAIDS and the proportion testing positive in my analysis, but much lower than those from Thembisa and Spectrum.

Table 9.2 - Estimates of the final rate of MTCT, by year of birth

Year of birth	Thembisa	Spectrum	UNAIDS	SAPMTCTE	Proportion of children tested with a positive PCR test
2010	11.06%	11.91%	9.00%		6.60%
2011	8.91%	12.28%	6.00%		4.60%
2012	7.71%	12.32%	4.50%		3.50%
2013	6.78%	10.88%	3.00%	4.30%	3.10%
2014	5.85%	10.37%	2.00%		2.50%
2015	5.02%	11.58%	2.00%		2.30%
2016	4.37%	9.85%	2.00%		1.00%

Note: The estimate from SAMPTCTE included infants born between October 2012 and May 2013.

Figure 9.1 - Estimates of the final rate of MTCT, by year of birth and data source
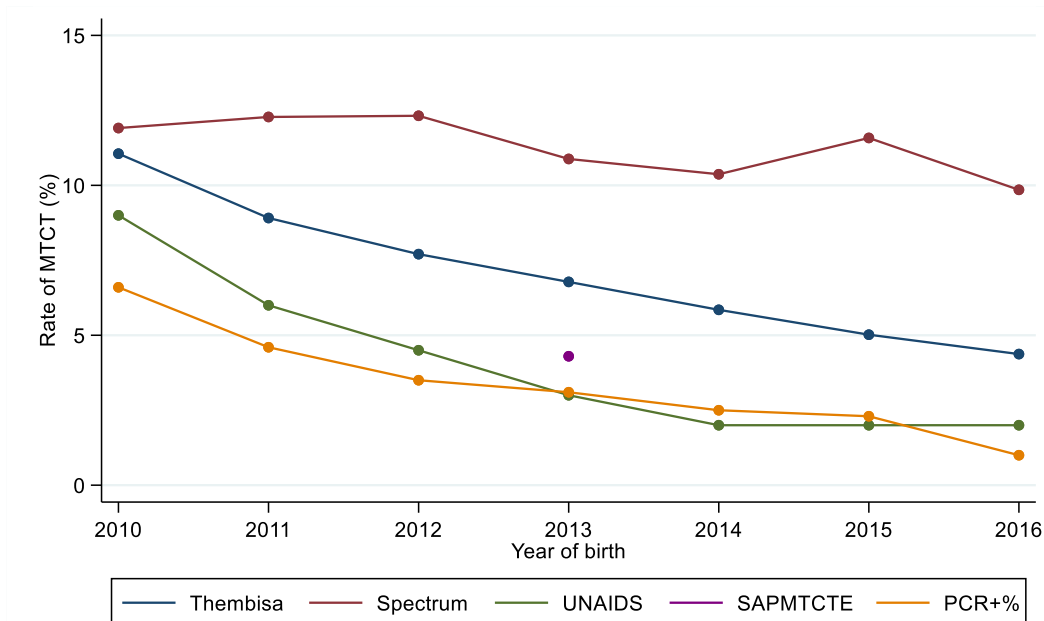
Note: The estimate from SAMPTCTE included infants born between October 2012 and May 2013.

For estimation of the number of children who acquired HIV but were not diagnosed, estimates of the number of children diagnosed with HIV were first required. The number of children born in the sub-district and diagnosed with HIV was 787 (Table 9.3), of whom 458 had a positive PCR test, and a further 329 were assumed diagnosed as they were on ART and assumed to have been born in sub-district either because they had a linked negative and/or indeterminate PCR test (n=9), or they initiated treatment within the sub-district (n=290), or both (n=30). An additional 106 children on ART in the sub-district were excluded because they transferred in after initiating

treatment elsewhere and had no linked PCR tests (and so were assumed to have been born outside of the sub-district).

Table 9.3 - Number of children diagnosed with HIV and assumed to have been born in the sub-district

Year of birth	Number of children with a positive PCR test	Number of children with no positive PCR test but on ART	Total number of children diagnosed with HIV
2010	69	56	125
2011	96	65	161
2012	78	42	120
2013	67	52	119
2014	60	49	109
2015	64	38	102
2016	24	27	51
Total	458	329	787

Table 9.4 shows the estimates of the total number of children with HIV (column C) and of the number who were not diagnosed (column E), based on the estimated rates of MTCT from Thembisa. Multiplying these rates by the number of HIV-exposed children born each year, 1,165 children born in the sub-district were estimated to have acquired HIV. Subtracting the 787 known to have been diagnosed, 378 (32%) were therefore estimated to have not been diagnosed. The proportion not diagnosed increased over time from 12% among those born in 2010 to 62% in 2016.

Table 9.4 - Estimates of the total number of children with HIV and the number of undiagnosed children, by year of birth, based on the estimated rate of MTCT from Thembisa

Year of birth	Transmission rate	Number of HIV-exposed children	Number of children with HIV	Number of diagnosed children	Number of undiagnosed children	Proportion of children undiagnosed
Source of data	Thembisa	Chapter 5, method 2		Table 9.3		
Calculation	A	B	C=A*B	D	E=C-D	F=E/C
2010	11.06%	1,280	141.52	125	16.52	12%
2011	8.91%	2,275	202.70	161	41.70	21%
2012	7.71%	2,400	184.92	120	64.92	35%
2013	6.78%	2,898	196.47	119	77.47	39%
2014	5.85%	2,620	153.23	109	44.23	29%
2015	5.02%	3,000	150.60	102	48.60	32%
2016	4.37%	3,097	135.43	51	84.43	62%
Total		17,570	1,165	787	378	32%

Note: Estimates for 2010 include those born from June onwards.

Based on these Thembisa estimates, a summary flowchart of the HIV-exposed children born within the sub-district is shown in Figure 9.2. Of the estimated 17,570 HIV-exposed children born in the sub-district between June 2010 and December 2016, 1,165 were estimated to have acquired HIV and 16,405 were estimated to have not. As described above, a total of 787 (68%) of the children with HIV were ever diagnosed in the sub-district during this time period, resulting in 378 (32%) of the remaining children with HIV being in the undiagnosed category. Of the 15,234 children with a PCR test from NHLS, 458 ever tested positive and 39 with only negative or indeterminate PCR tests were also assumed to have HIV as they started ART. This means the remaining 14,737 (= 15,234 - 458 - 39) who received a PCR test represent 90% of the estimated HIV-negative children, leaving 1,668 (= 16,405 - 14,737) (10%) HIV-negative children who did not receive a PCR test.

Figure 9.2 - Flowchart of HIV-exposed children born in the sub-district

Estimates of the proportion of children with HIV who were not diagnosed based on the rates of transmission from Spectrum and UNAIDS are shown in Table 9.5 and Table 9.6 respectively. As the estimates of the rate of MTCT from Spectrum were higher than Thembisa, the corresponding estimate of the proportion of children not diagnosed was also higher at 60% across the whole time period. Using the UNAIDS estimates, the overall proportion of missed infections was 4%, although for some calendar years estimates of the total number of children with HIV were lower than the number diagnosed, resulting in a negative number of undiagnosed children.
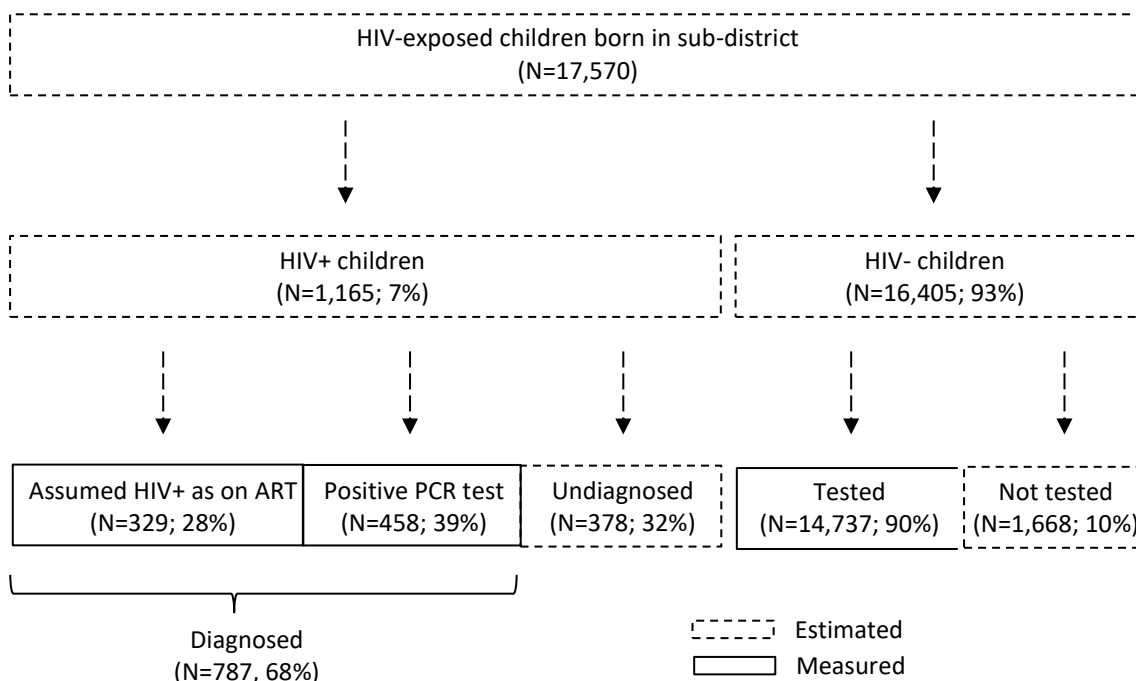
Table 9.5 - Estimates of the total number of children with HIV and the number of undiagnosed children, by year of birth, based on the estimated rate of MTCT from Spectrum

Year of birth	Transmission rate	Number of HIV-exposed children	Number of children with HIV	Number of diagnosed children	Number of undiagnosed children	Proportion of children undiagnosed
Source of data	Spectrum	Chapter 5, method 2		Table 9.3		
Calculation	A	B	C=A*B	D	E=C-D	F=E/C
2010	11.91%	1,280	152.45	125	27.45	18%
2011	12.28%	2,275	279.37	161	118.37	42%
2012	12.32%	2,400	295.68	120	175.68	59%
2013	10.88%	2,898	315.30	119	196.30	62%
2014	10.37%	2,620	271.69	109	162.69	60%
2015	11.58%	3,000	347.40	102	245.40	71%
2016	9.85%	3,097	305.05	51	254.05	83%
Total		17,570	1,967	787	1,180	60%

Note: Estimates for 2010 include those born from June onwards.

Table 9.6 - Estimates of the total number of children with HIV and the number of undiagnosed children, by year of birth, based on the estimated rate of MTCT from UNAIDS

Year of birth	Transmission rate	Number of HIV-exposed children	Number of children with HIV	Number of diagnosed children	Number of undiagnosed children	Proportion of children undiagnosed
Source of data	UNAIDS	Chapter 5, method 2		Table 9.3		
Calculation	A	B	C=A*B	D	E=C-D	F=E/C
2010	9.00%	1,280	115.20	125	-9.80	-9%
2011	6.00%	2,275	159.25	161	-1.75	-1%
2012	4.50%	2,400	144.00	120	24.00	17%
2013	3.00%	2,898	144.90	119	25.90	18%
2014	2.00%	2,620	104.80	109	-4.20	-4%
2015	2.00%	3,000	90.00	102	-12.00	-11%
2016	2.00%	3,097	61.94	51	10.94	18%
Total		17,570	820	787	33.00	4%

Notes: The proportion of children undiagnosed is less than 0% for some calendar years, as the estimated number of children with HIV is lower than the number diagnosed. Estimates for 2010 include those born from June onwards.

9.4.2. The cascade of care

The estimates of cascade #1 are shown in Table 9.7. Of the estimated 1,165 children born in the sub-district across the whole time period who acquired HIV, 51 had died and 80 were known to have transferred to a clinic outside of the sub-district by the 30th June 2017, leaving a denominator of 1,034 children alive and still living in the sub-district. Of these, 660 (64%) had been diagnosed, of whom 197 (30%) were on ART on the 30th June 2017. Of those on ART, 26 had an available viral load measurement, of whom 18 (69%) were suppressed ≤400 copies/mL. Assuming the same proportion of children on ART but with no viral load measurement available were suppressed, an estimated 13% of the total denominator of 1,034 children were suppressed. The results from the cascade are shown graphically in Figure 9.3 (where percentages are calculated using the denominator of the number meeting the previous stage) and Figure 9.4 (where percentages are calculated using the denominator of the total number of children with HIV living in the sub-district, indicated by the red bar, for each stage), with the 90-90-90 targets shown by the coloured bars.

Table 9.7 - Cascade #1: Current status of children alive and living with HIV in the sub-district as of 30th June 2017

Stage	n	% of those meeting the previous stage	% overall
Born in sub-district who acquired HIV	1,165		
No longer in sub-district on 30th June 2017	131		
Died	51		
Transferred out	80		
Still in sub-district on 30th June 2017	1,034		
Diagnosed	660	64% (660/1,034)	64% (660/1,034)
On ART	197	30% (197/660)	19% (197/1,034)
On ART and with a viral load available	26		
Virally suppressed	18	69% (18/26)	13% ((69% x 197)/1,034)*

*Calculated assuming the same proportion of children were suppressed among those without a viral load as observed among those with a viral load

220

Figure 9.3 - Cascade #1: Current status of children alive and living with HIV in the sub-district as of 30th June 2017 (percentages of those meeting the previous stage)
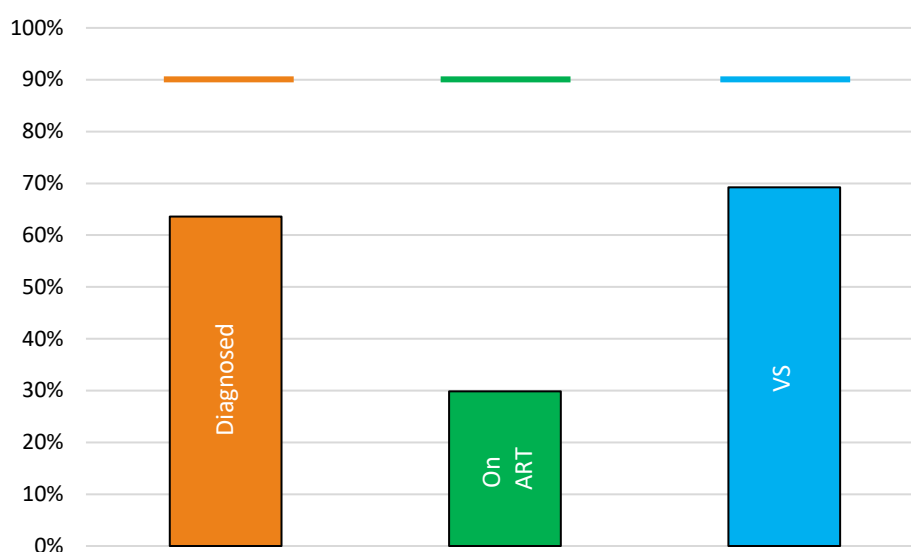
Notes: Bars indicate 90-90-90 targets. VS: virally suppressed. Denominator for each stage is the number meeting the previous stage, and thus the total number with HIV is not shown.

Figure 9.4 - Cascade #1: Current status of children alive and living with HIV in the sub-district as of 30th June 2017 (overall percentages)
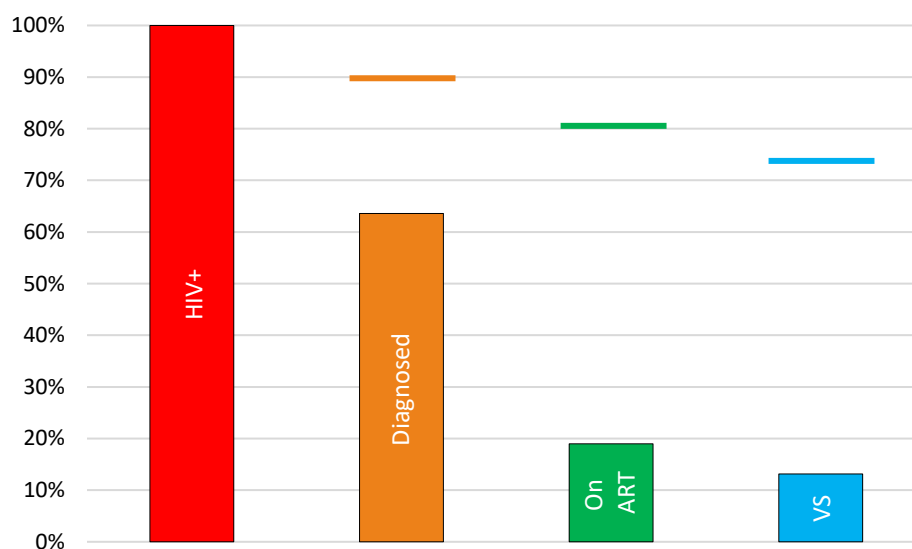
Notes: Bars indicate 90-90-90 targets. VS: virally suppressed. Denominator for each stage is the number of children with HIV (the red bar).

The results of cascade #2 are shown in Table 9.8. Of 12,973 HIV-exposed children born across the sub-district by 30th June 2015, an estimated 955 (7%) acquired HIV and 12,018 (93%) did not. In total, 11,433 (88%) HIV-exposed children received a PCR test by 2 years of age, consisting of 561 (59%) of the children who had HIV and 10,872 (90%) of the children who did not have HIV. Less than 1% of all HIV-exposed children were known to have died before receiving a PCR test. Of the 561 children with HIV who were diagnosed, 362 (65%) initiated ART by 2 years of age, corresponding to 38% of all the children with HIV. Of the 362 who initiated ART, 192 (53%) were

still on ART at 2 years of age (corresponding to 20% of the estimated total number of children with HIV), 21 (6%) were still in follow-up but no longer on ART, 29 (8%) had transferred to a clinic outside of the sub-district, 28 (8%) had died, 55 (15%) were lost-to-follow-up after ART initiation, and the follow-up status of the remaining 35 (10%) was unknown. Of the 192 children still on ART, 100 (52%) had no viral load measurement available. Of the remaining 92, 63 (68%) were virally suppressed ≤400 copies/mL and 29 (32%) were not. Assuming the same proportion of children on ART with no viral load measurement available were suppressed as among those with a measurement, this corresponds to 6% of the total 955 children with HIV being suppressed. The results are shown graphically in Figure 9.5, where the denominator for the first two bars is the number of HIV-exposed children, for the final three is the number of children with HIV.

Table 9.8 - Cascade #2: Status at 2 years of age of all children born in the sub-district and exposed to HIV

	Number of children	% of those meeting the previous stage	% overall**
HIV exposed born	12,973		
HIV- born	12,018		
HIV+ born	955		
HIV exposed tested	11,433	88%	88%
HIV exposed died before tested	49	<1%	<1%
HIV exposed alive but not tested	1,444	11%	11%
HIV- tested	10,872	90%	90%
HIV- died before tested	45	<1%	<1%
HIV- alive but not tested	1,101	9%	9%
HIV+ diagnosed	561	59%	59%
HIV+ died before diagnosis	4	<1%	<1%
HIV+ alive but not diagnosed	390	41%	41%
Initiated ART	362	65%	38%
Died before initiating ART	10	2%	1%
Did not initiate ART	189	34%	20%
Currently on ART	192	53%	20%
Currently off ART	21	6%	2%
Transferred out	29	8%	3%
Died	28	8%	3%
Lost-to-follow-up	55	15%	6%
Unknown	35	10%	4%
Virally suppressed***	63	68%	14%
Not virally suppressed***	29	32%	6%
No viral load available	100		

Includes those born to 30th June 2015; ** Denominator for the overall percentages is the number of children with the respective HIV status for the testing stage, and the number of children with HIV for all subsequent stages; *** For the proportion virally suppressed ≤400 copies/mL, the percentage of those mee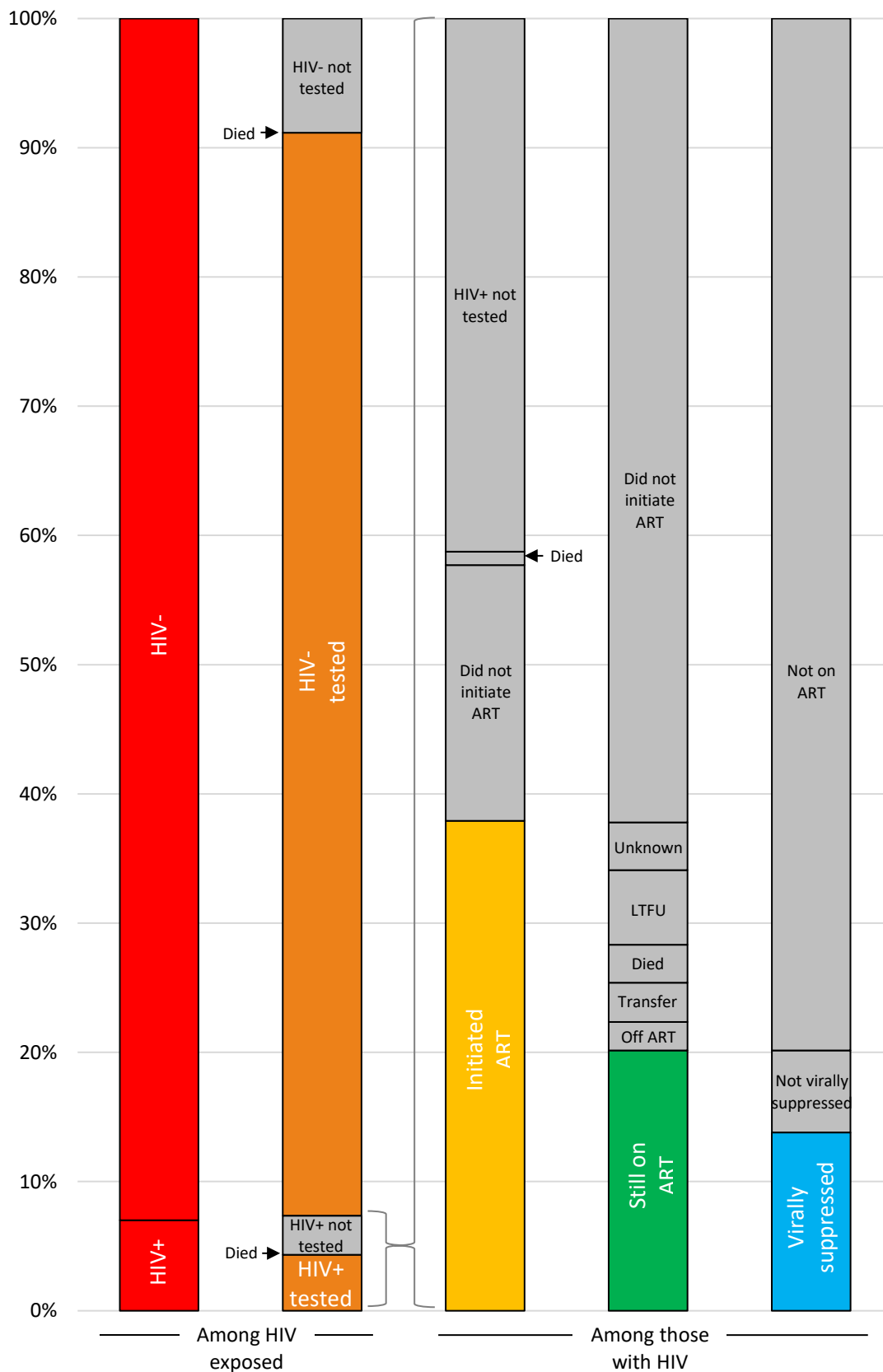ting the previous stage is calculated based on those with an available measurement, and the overall percentage is calculated assuming the same proportion of children on ART with no viral load measurement available were suppressed as among those with a measurement.

Figure 9.5 - Cascade #2: Status at 2 years of age of all HIV-exposed children born in the sub-district (overall percentages)

Notes: Includes those born to 30th June 2015. The denominator is the number of HIV-exposed children with the respective HIV status for the testing stage, and the number of children with HIV for all subsequent stages. For the proportion virally suppressed, the percentage is calculated assuming the same proportion of children on ART with no viral load measurement available were suppressed as among those with a measurement. LTFU: Lost-to-follow-up.
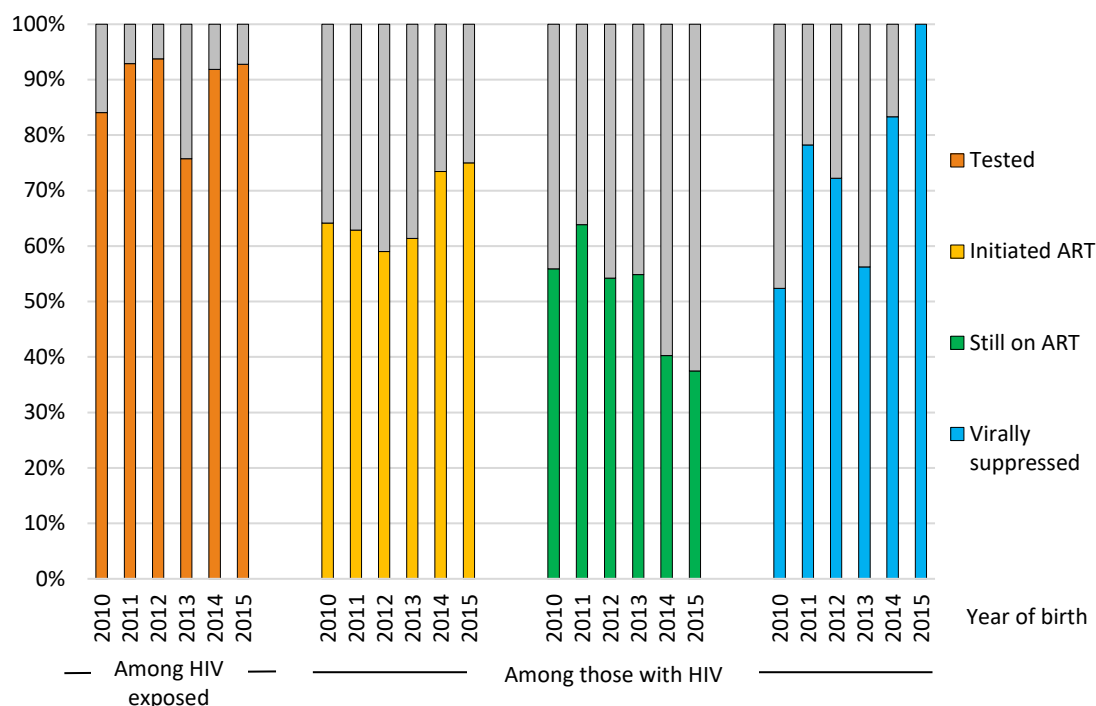
Table 9.9 - Cascade #2: Status at 2 years of age of all HIV-exposed children born in the sub-district, by year of birth

Stage	2010			2011			2012			2013			2014			2015*		
	Number of children	% of those meeting the previous stage	% overall**	Number of children	% of those meeting the previous stage	% overall**	Number of children	% of those meeting the previous stage	% overall**	Number of children	% of those meeting the previous stage	% overall**	Number of children	% of those meeting the previous stage	% overall**	Number of children	% of those meeting the previous stage	% overall**
HIV exposed born	1,280			2,275			2,400			2,898			2,620			1,500		
HIV- born	1,138			2,072			2,215			2,702			2,467			1,424		
HIV+ born	142			203			185			196			153			76		
HIV exposed tested	1,076	84%	84%	2,113	93%	93%	2,251	94%	94%	2,195	76%	76%	2,406	92%	92%	1,392	93%	93%
HIV exposed died before tested	4	<1%	<1%	12	1%	1%	14	1%	1%	6	<1%	<1%	8	<1%	<1%	5	<1%	<1%
HIV exposed alive but not tested	196	15%	15%	138	6%	6%	121	5%	5%	693	24%	24%	198	8%	8%	98	7%	7%
HIV- tested	970	76%	76%	1,981	87%	87%	2,151	90%	90%	2,094	72%	72%	2,308	88%	88%	1,368	91%	91%
HIV- died before tested	4	<1%	<1%	11	<1%	<1%	13	1%	1%	6	<1%	<1%	7	<1%	<1%	4	<1%	<1%
HIV- alive but not tested	164	13%	13%	80	4%	4%	51	2%	2%	602	21%	21%	152	6%	6%	52	3%	3%
HIV+ diagnosed	106	8%	8%	132	6%	6%	100	4%	4%	101	3%	3%	98	4%	4%	24	2%	2%
HIV+ died before diagnosis	0	<1%	<1%	1	<1%	<1%	1	<1%	<1%	0	<1%	<1%	1	<1%	<1%	1	<1%	<1%
HIV+ alive but not diagnosed	36	3%	3%	70	3%	3%	84	4%	4%	95	3%	3%	54	2%	2%	51	3%	3%
Initiated ART	68	64%	48%	83	63%	41%	59	59%	32%	62	61%	32%	72	73%	47%	18	75%	24%
Died before initiating ART	0	0%	0%	2	2%	1%	4	4%	2%	4	4%	2%	0	0%	0%	0	0%	0%
Did not initiate ART	38	36%	27%	47	36%	23%	37	37%	20%	35	35%	18%	26	27%	17%	6	25%	8%
Currently on ART	38	56%	27%	53	64%	26%	32	54%	17%	34	55%	17%	29	40%	19%	6	33%	8%
Currently off ART	2	3%	1%	3	4%	1%	5	8%	3%	6	10%	3%	4	6%	3%	1	6%	1%
Transferred out	2	3%	1%	2	2%	1%	5	8%	3%	7	11%	4%	11	15%	7%	2	11%	3%
Died	10	15%	7%	11	13%	5%	1	2%	1%	2	3%	1%	2	3%	1%	2	11%	3%
Lost-to-follow-up	7	10%	5%	9	11%	4%	14	24%	8%	8	13%	4%	13	18%	8%	4	22%	5%
Unknown	9	13%	6%	5	6%	2%	2	3%	1%	5	8%	3%	13	18%	8%	1	6%	1%
Virally suppressed***	11	52%	14%	18	78%	20%	13	72%	12%	9	56%	10%	10	83%	16%	2	100%	8%
Not virally suppressed***	10	48%	13%	5	22%	6%	5	28%	5%	7	44%	8%	2	17%	3%	0	0%	0%
No viral load available	17			30			14			18			17			4		

* Includes those born to 30th June 2015; ** Denominator for the overall percentages is the number of children with the respective HIV status for the testing stage, and the number of children with HIV for all subsequent stages; *** For the proportion virally suppressed, the percentage of those meeting the previous stage is calculated based on those with an available measurement, and the overall percentage is calculated assuming the same proportion of children on ART with no viral load measurement available were suppressed as among those with a measurement

Figure 9.6 - Cascade #2: Status at 2 years of age of all HIV-exposed children born in the sub-district, by year of birth (percentages of those meeting the previous stage)

Notes: Includes those born to 30th June 2015. The denominator is the number of HIV-exposed children for the testing stage, the number of children diagnosed with HIV for the ART initiation stage, the number of children who initiated ART for the still on ART stage, and the number on ART and with an available measurement for the viral suppression stage.

Figure 9.7 - Cascade #2: Status at 2 years of age of all HIV-exposed children born in the sub-district, by year of birth (overall percentages)

Notes: Includes those born to 30th June 2015. The denominator is the number of children with the respective HIV status for the testing stage, and the number of children with HIV for all subsequent stages. For the proportion virally suppressed, the percentage is calculated assuming the same proportion of children on ART with no viral load measurement available were suppressed as among those with a measurement.
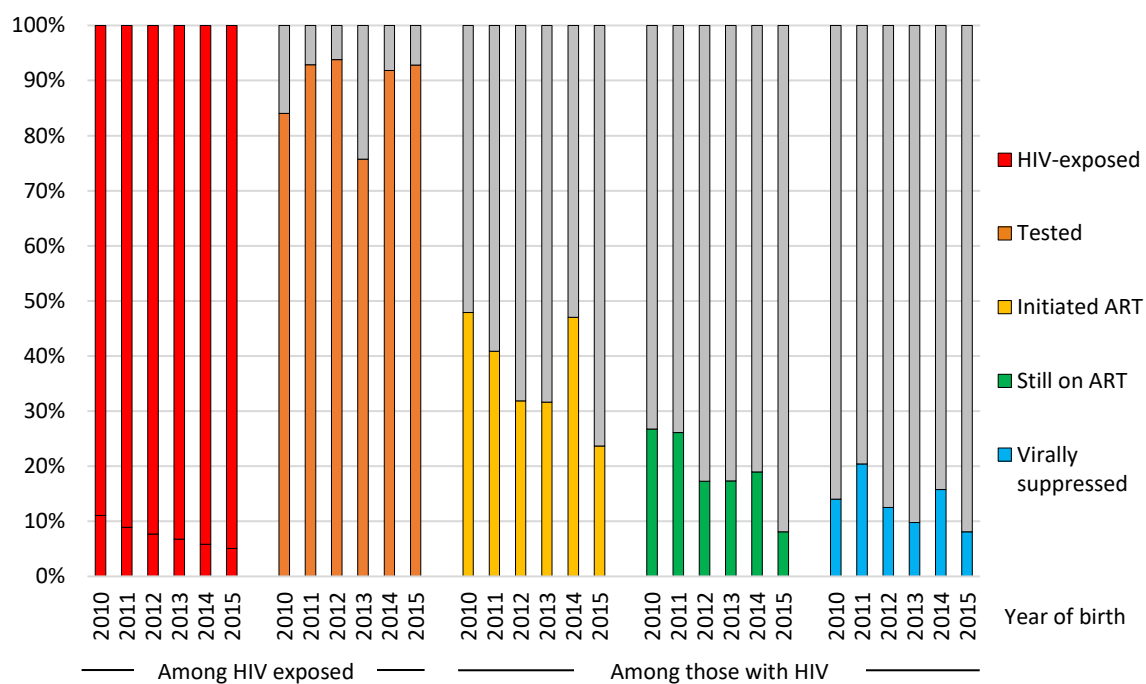
The results of cascade #2 are shown by year of birth in Table 9.9, Figure 9.6 and Figure 9.7. Over time, the total proportion of HIV-exposed children tested was relatively steady over time at between 84% and 93%, although was lower for those born in 2013 at 76%. In all years, a lower proportion of those with HIV were tested compared to those who did not have HIV, with the difference between the two groups increasing for those born in later years. Among the children with HIV, the proportion of those diagnosed who went on to initiate ART by 2 years of age was between 59% and 61% each year for those born between 2010 and 2013, and subsequently increased to 73% and 75% of those born in 2014 and 2015 respectively. Conversely, while the proportion of these children who were retained on ART by 2 years was relatively constant at between 54% and 64% until 2013, it subsequently dropped to 40% in 2014 and 33% in 2015. The proportion who transferred out after ART initiation increased over time (from 3% among those born in 2010 to 11% among those born in 2015), as did the proportion lost-to-follow-up (from 10% to 22%). The proportion reported to have died on ART decreased from 15% and 11% among those born in 2010 and 2011, to between 2% and 3% among those born in 2012, 2013 and 2014, although subsequently increased again to 11% of those born in 2015 (which may be the result of the smaller denominator for this year). There was no real trend in the proportion currently off ART or the proportion with unknown follow-up status. Among those reported to be on ART at 2 years of age, the proportion virally suppressed among those with a viral load measurement available, varied substantially over time at between 52% and 100%. Overall, the proportion of all children with HIV who were virally suppressed (assuming the same proportion suppressed among those with no measurement) at age 2 years varied over time, although with no discernible trend, at between 8% and 20%.

9.5. Discussion

Three different sets of estimates of the rate of MTCT were used to estimate the total number of children born in the sub-district who acquired HIV, with high variation in the resulting estimates of the number and proportion of children who were not diagnosed. Using the rates from Thembisa, about one-third of children with HIV were not diagnosed, and this proportion generally increased over time, from 12% among those born in 2010 to 32% among those born in 2015, with a larger increase to 62% of children born in 2016. Increased PMTCT coverage has led to a higher proportion of children acquiring HIV doing so in the postpartum period rather than in utero or intrapartum, with, given the low coverage of repeat testing described in Chapter 7, these children less likely to be diagnosed; this may explain why the proportion not diagnosed increased over time. It is also possible that the true rate of MTCT actually fell faster than estimated by the Thembisa model, resulting in the overestimation of the number of children who acquired HIV in later years and thus the overestimation of the proportion not diagnosed. The analysis based on data from UNAIDS resulted in fewer children estimated to have acquired

HIV than were diagnosed (that is an impossible negative number of children were not diagnosed), suggesting the rates of MTCT (which were 2% between 2014 and 2016) were too low. Estimates of the rate of MTCT from Spectrum were the highest and did not decrease over time, which seems unlikely.

The advantages and disadvantages of the methods used for each of the two cascades reported are described below, followed by a summary of the results, and of the issues and limitations of the estimation of each cascade stage, as well as the likely direction of the effect of the resulting bias on the estimate of cascade achievement (Table 9.10).

Cascade #1 was designed to assess achievement of the UNAIDS 90-90-90 targets within the sub-district, summarising the status of those who acquired HIV and were still alive and living within the sub-district on the 30th June 2017. Cascades designed in this way are useful to give a snapshot of the status of a population of interest at a given point in time. However, because this thesis only included those born since June 2010 and followed all individuals from birth, the maximum follow-up time available increased over calendar time, meaning interpreting results from cascades assessed in this way at repeated time points would have been difficult.

Cascade #2 was designed to assess the performance of the sub-district's early infant diagnosis program, summarising the status at 2 years of age of all children exposed to HIV. An advantage of this method is that it allowed the quantification of mortality and loss-to-follow-up, with these children simply excluded from the denominator of cascade #1. Conversely, this method also has several disadvantages. Firstly, some individuals were known to have transferred out of the sub-district, many of whom may actually be in care elsewhere (with some others classified here as lost-to-follow-up likely to have transferred as well); this cascade can only summarise HIV care within sub-district. Secondly, only those born prior to June 2015 could be included, with those born later not having enough follow-up time. Finally, as currently conducted, the cascade does not show what happens to children past two years of age, although the analysis could potentially be repeated at older ages using the same methods.

By two years of age, 88% of HIV-exposed children born in the sub-district had been tested, and of those children diagnosed with HIV, 65% had initiated ART, of whom 53% were still on ART, of whom 68% were virally suppressed. Among the 1,034 children born between June 2010 and December 2016 and still living with HIV in the sub-district on the 30th June 2017, 64% were estimated to have been diagnosed, of whom 30% were estimated to be on ART, of whom 69% were estimated to be virally suppressed. Looking over calendar time, there were few clear trends, possibly because of high uncertainty and limitations in the estimation of all stages, as described below.

Cascade #2 demonstrated that a lower proportion of children who had acquired HIV were ever tested than of those who did not acquire HIV. This result may be expected, with a higher risk of transmission known to be associated with poor maternal engagement in care, which is known to be associated with poor infant engagement in care [220]. The proportion of children with HIV who were tested decreased over calendar time, although there was no clear trend in testing coverage among children without HIV.

As well as potential inaccurate estimation of the rate of MTCT by Thembisa, there are several other causes of uncertainty around the estimation of testing coverage. Firstly, there are several sources of uncertainty around the estimate of the number of HIV-exposed children as discussed in Chapter 5, including the antenatal seroprevalence estimate being based on small numbers and having to be adjusted for incident HIV during pregnancy, and the underestimation of births due to late registration. Secondly, for the purposes of calculating the proportion of children diagnosed by a specific time point, it was assumed that all children had acquired HIV in utero or intrapartum, thus ignoring the fact that some children would not yet have acquired HIV (and thus could not yet be diagnosed). This may have led to the underestimation of the true proportion of children with HIV who were diagnosed. Thirdly, only diagnoses based on PCR tests, typically conducted only up to approximately 18 months of age, were accounted for (unless the child subsequently initiated ART), as no data on antibody testing were available, leading to underestimation of coverage. Fourthly, limitations of the NHLS data may have resulted in the incorrect linkage of repeat PCR tests; although underlinkage (and thus overestimation of testing) was more likely, this may have biased the estimate in either direction. Finally, limitations of data from both NHLS and ACCDB may have led to the failure to link all children with a positive PCR test to their ART data; these individuals would therefore have been counted as two diagnosed children, one with a PCR test but not on ART and one on ART with no PCR test, leading to the overestimation of testing.

In both cascades, initiation of and retention on ART was shown to be the main bottleneck to completion of the cascade, with the proportion meeting this stage much lower than the others. There are however several reasons why achievement of this stage may have been underestimated. Firstly, although data on death and migration were available through ACCDB for those on ART, prior to ART initiation there were no data on migration and data on deaths were only available for the relatively small subset who were linked to demographic surveillance. This would have resulted in children no longer residing in the sub-district incorrectly being included in denominator for cascade #1 and in misclassification of children in cascade #2. Secondly, the date of ART initiation was unknown for children who transferred into the sub-district already on treatment (but who were assumed to have been born in the sub-district because they had received a PCR test there), with the (later) date of transfer used instead,

228

resulting in the underestimation of the proportion on ART at any given time point. Thirdly, the date of ART initiation was unknown and therefore imputed for a small number of children for whom no ART follow-up was recorded in ACCDB (as described in Chapter 8), which may have led to the incorrect classification (in either direction) of their ART status at each time point. Finally, errors in the linkage conducted by AHRI when data from all the clinics are pooled (to identify patients moving between clinics in the sub-district) would have resulted in their incorrect classification as having transferred out of the sub-district.

The analysis of viral suppression may have been biased because of the high proportion of children with no viral load data available. Although it was assumed that data were missing at random, either because of incomplete recording in ACCDB or because of low rates of monitoring, it is possible that viral load monitoring was targeted at those suspected of virological failure, which would lead to the underestimation of the proportion who achieved viral suppression.

Although the limitations discussed result in uncertainty over the accuracy of the actual estimates made, the relative trend in achievement of each of the stages looks consistent with other South African studies. In the sub-district specifically, linkage to care and initiation of treatment have been shown to be poor in those aged over 15 years [102, 273]. Similarly, in 2017, UNAIDS estimated South Africa's progress towards the achievement of the 90-90-90 targets to be 78-74-78 among children aged under 15 years, and 90-68-77 across the population as a whole [2]. Only one other study has assessed the complete cascade among infants in South Africa [177]; among 838 infants exposed to HIV and born at a single hospital in Johannesburg between August 2008 and July 2010, 82% received a PCR test at 6 weeks of age, and of those diagnosed with HIV, 61% started ART, of whom 54% achieved viral suppression in a median 18 weeks.

9.6. Key findings

The key findings from this chapter are:

- By two years of age, 88% of HIV-exposed children born in the sub-district had been tested, and of those children who were diagnosed with HIV, 65% had initiated ART, of whom 53% were still on ART, of whom 68% were virally suppressed.
- Among the 1,034 children born between June 2010 and December 2016 and still living with HIV in the sub-district on the 30th June 2017, 64% were estimated to have been diagnosed, of whom 30% were estimated to be on ART, of whom 69% were estimated to be virally suppressed.

Table 9.10 - Sources of uncertainty across cascade

Stage	Issue	Type of error	Likely direction of effect of bias on estimate of cascade
Number of HIV-exposed children	Seroprevalence based on small numbers by district	Reliability	Either
	Women acquiring HIV during pregnancy but after sampling not identified by ANCHSS (although adjustment made)	Misclassification	Overestimation
	Underestimation of births due to late registration (although adjustment made)	Misclassification	Overestimation
	Data not directly linked to subsequent stages	Validity	Either
Number of children with HIV	Uncertainty in the estimated rate of MTCT from Thembisa	Modelling error	Either
Number of children tested/ diagnosed	Assuming all children born with HIV, but may not have acquired HIV by time point of analysis so couldn't be expected to be diagnosed	Misclassification	Underestimation
	Only know about those diagnosed up to 18 months of age (i.e. with a PCR test)	Missing data	Underestimation
	NHLS data limitations may have led to failure to identify repeat tests on the same children	Misclassification	Overestimation
	Tests from different children may have been incorrectly linked	Misclassification	Underestimation
	Assumptions made when defining which children were born in sub-district may have been incorrect	Misclassification	Either
	No data on migration, and data on deaths were only available for the small subset who were linked to ACDIS	Misclassification	Underestimation
	Poor data quality causes difficulty linking NHLS to ACCDB, meaning some children with a PCR test may not have been linked to their ART data, leading to their double counting in the diagnosis stage (once with a PCR test but not on ART, and once on ART with no PCR test)	Linkage error	Underestimation
Number of children with HIV who initiated ART	No data on migration, and data on deaths were only available for the small subset who were linked to ACDIS	Misclassification	Underestimation
	Some individuals (who were assumed to have been born in the sub-district as they had a linked PCR test) initiated ART outside the sub-district and later transferred in; their date of ART initiation was unknown so (later) date first seen on ART in sub-district used instead	Misclassification	Underestimation
	Some individuals on ART had an unknown start date, imputed based on average age at initiation for year of birth	Misclassification	Either
Number of children with HIV currently on ART	Failure to link records in TIER.net for individuals who transferred between clinics would result in incorrect classification as having transferred out of the sub-district	Misclassification	Underestimation
	Many individuals who transferred to a clinic outside of the sub-district are likely still on ART	Misclassification	Underestimation
Number of children on ART who achieved viral suppression	Viral load data was assumed to be missing at random, but is possible that testing was targeted at those suspected of virological failure	Selection bias	Underestimation

Chapter 10. Discussion

10.1. Introduction

Here I outline the key findings from each chapter, describe the main strengths and limitations of this thesis, and make recommendations both for future work to explore the cascade and for improvements to data collection.

10.2. Summary and relevance of key findings

In Chapter 4, I conducted data linkage within the NHLS PCR dataset to identify repeat tests conducted on the same infant, and subsequently between infants with a PCR test and each of ACCDB (the ART clinical database) and ACDIS (demographic surveillance), enabling estimation of the cascade in the rest of the thesis. This is the first time data from NHLS and TIER.net (on which ACCDB is based) have been linked to follow infants diagnosed with HIV in South Africa longitudinally. A number of limitations of the data available within the NHLS dataset were identified, including the poor recording of South African national ID and the use of mother's name instead of the infant's, which may have had an impact on results throughout the thesis. No gold standard dataset of linked records existed to fully evaluate the quality of any of the linkage steps, however, the proportion of infants in the deduplicated NHLS database with the same surname, date of birth and sex as another infant was compared to the proportion that would be expected due to chance agreement, which suggested some underlinkage.

In Chapter 5, I compared estimates of the proportion of HIV-exposed infants who received a PCR test by 7 weeks of age and the proportion who were ever tested, using four different methods. For the proportion ever tested, the first method was based on linking NHLS PCR data to demographic surveillance data from ACDIS. These estimates were the lowest of the four, at between 29% and 48% by year of birth, with the main limitations being likely underlinkage and data being missing from one clinic for part of the time period. The second method was based on the number of infants in the deduplicated NHLS PCR dataset, divided by an estimate of the number of HIV-exposed infants calculated from the number of live births (from SSA) and the antenatal seroprevalence (from ANCHSS). The proportion of infants receiving a PCR test here was between 75% and 99% per year. Adjustments were made to account for late reporting of live births and of incident HIV during pregnancy not captured by ANCHSS, although the completeness of these adjustments is unclear. The third approach calculated coverage as the number of infants with an NHLS PCR test divided by an estimate of the number of HIV-exposed infants born as reported to the Department of Health through DHIS. These estimates were impossibly greater than 100%, at between 101% and 160% by year of birth. The final method

used data from Road-to-health booklets (collected in the MONARCH trial), which could only be used to estimate the proportion tested by their 6 week post-natal visit. Estimates were somewhere in the middle of those calculated for this same outcome with the other methods, although there was some evidence of incomplete recording of the tests in the booklets, particularly towards the end of the trial.

The strengths of this analysis were the variety of data sources used, including both individual-level and aggregated data, data from patient records and data from clinic reports. I considered limitations to the sources of data used for each method and the likely impact on the resulting estimate of testing coverage, making adjustments to estimates to overcome these where possible, with the estimates made on the unadjusted data very different. One previous study compared estimates made using two methods (comparable to methods 2 and 3 used here) and explored the effect of different definitions of coverage based on the DHIS data [123]. However, this earlier study did not consider all limitations, for example, the authors did not deduplicate infants with repeat PCR tests within the NHLS dataset.

In Chapter 6, I considered which maternal, pregnancy and delivery related factors were associated with PCR testing, using data from both NHLS-ACDIS linkage and from Road-to-health booklets as collected for the MONARCH trial. Although there was high variation in the estimates made in Chapter 5, coverage was unlikely to be 100% and it is therefore important to identify the infants most likely to be missed. The differences in the estimated coverage between the two methods suggest some misclassification of the outcome in each dataset, which would lead to the attenuation of any true effects of the covariates investigated. As a result, few of the covariates investigated showed any significant effect, although infants born to women with a higher number of previous live-born children and with a longer time since ART initiation, who are likely to be better engaged in care and may have a greater understanding of the potential for transmission and the need for infant testing, were more likely to be tested. These results were consistent with those observed in other studies, both in South Africa and elsewhere [242, 244, 274]. There was a high proportion of missing data across some of the covariates of interest, such as maternal viral load at delivery, which led to their exclusion from the multivariable models, and no data were available on some other covariates which may have been important predictors, such as maternal HIV status disclosure (e.g. to partners and/or family). Conflicting trends over time were observed, likely the result of the limitations of each method.

In Chapter 7, I considered the frequency, timing and results of PCR testing, among those infants who received a test, in order to assess adherence to the national testing guidelines, particularly after the introduction of birth testing in April 2015. Less than a quarter of infants testing negative at birth ever received another PCR test, leading to the possibility of those who acquired HIV late in pregnancy, at delivery, or in the postpartum breastfeeding period not being diagnosed.

Previous research from South Africa assessing this has focussed on healthcare facilities in urban areas (where repeat testing has been estimated to be higher at between 70% and 80%) [160, 164, 165], or has been compromised by methodological difficulties in identifying repeat PCR tests in the NHLS dataset [122]. Healthcare facilities in rural areas are often less well resourced, and are less likely to be academic hospitals or centres of excellence, meaning retention through the cascade is likely to be poorer. Many of the published studies described were from the same few clinics and hospitals; healthcare workers who conduct research may be more familiar with guidelines, limiting the generalisability of these results to other areas. As well as being important for understanding the success of birth testing in South Africa, these results are important for other high HIV burden and low- and middle-income countries, such as Kenya and Botswana [275, 276], which are considering its implementation. Similarly, confirmatory testing among infants with a positive test result was low, which may mean that infants who did not have HIV were initiated on ART, although there was some variation by clinic.

Vaccination coverage among HIV-exposed infants in the sub-district was high in the first year of life, which may suggest that suboptimal PCR testing was related to poor HIV care processes rather than poor engagement in broader child health (IMCI, integrated management of childhood illness) services. Although the two services are delivered in same health clinic, different nurses are responsible for each, and so HIV-exposed infants in need of PCR testing may not be being adequately identified or some may have sought PCR testing outside of their area of residence. Previous research from South Africa has demonstrated the challenges for healthcare workers trained in IMCI in identifying children with HIV, including concerns about asking caregivers about maternal HIV status, and inadequate training leading to poor understanding of the HIV component of IMCI [277]. Repeat testing coverage varied by clinic, which will provide an opportunity to identify the more successful processes. Vaccination data was reported by caregivers, with Road-to-health booklets seen for only approximately 10% of infants, and so social desirability bias may have led to the overestimation of coverage.

The proportion of infants tested who had a positive PCR test decreased over time to 1% among those born in 2016, though as expected was lower than the estimates of the true rate of MTCT from other sources summarised in Chapter 9, given the low proportion with a repeat test after birth and that some HIV-exposed infants were never tested at all. There was high variation in estimates of the rate of MTCT, however, even taking the lowest estimate of 2% in 2016, the sub-district is a long way off meeting the WHO targets for elimination of MTCT [40]; regions must meet the target of ≤50 new paediatric infections per 100,000 live births, which (given the high antenatal seroprevalence of over 45%) corresponds to a transmission rate of ≤0.1% in this setting.

In Chapter 8, outcomes among children diagnosed with HIV were considered, including the proportion who went on to initiate ART and subsequently achieve viral suppression. The proportion of infants with a positive PCR test result who started ART was estimated to be only 53% overall, though those born more recently were more likely to start. There was a high number of individuals on ART who had no linked PCR test; many of these were likely to have been diagnosed after 18 months of age with an antibody test rather than a PCR test (as they were older at treatment initiation) or diagnosed outside of the sub-district (as they initiated treatment outside the sub-district), but the remaining individuals may represent missed links between the NHLS PCR data and ACCDB. An upper bound on the estimate of proportion starting ART was thus calculated to be 77%. No other methods of estimating linkage to ART were available for comparison. Only one other study, conducted in 2012, has looked at linkage to ART in a rural setting in South Africa, reporting 56% of infants with a positive PCR test began treatment [175]. A high proportion of the infants who did start treatment in the sub-district presented late to care, with the two-thirds classified as CDC immunological stage 3 (corresponding to a CD4 count <750 cells/mm^3 or CD4% <26% for those aged less than 1 year for example) at initiation. There was a slight improvement over time, however even among those initiating in 2016 this proportion was still 61%; this is concerning given the evidence of the benefits of ART even at high CD4 counts.

Although linkage to ART improved in more recent calendar years, among those born after the introduction of birth testing, there was evidence that infants diagnosed at birth were less likely to go on to initiate ART than those diagnosed at an older age. One explanation for this could be that those diagnosed outside of routine PCR testing would be more likely to present symptomatically and thus be more likely to start treatment, although recording of the infant's mother's name on PCR test forms may occur more frequently at birth, making these tests harder to link to ART data, and thus leading to underestimation of ART initiation among those diagnosed at birth. This analysis was limited by the relatively small number of infants who acquired HIV following the change in the guidelines. Excluding studies assessing targeted birth testing [165, 184], for which results might be expected to be very different, only one other study has looked at the impact of birth testing on ART initiation; 96% of infants diagnosed at birth at a hospital in Johannesburg went on to start treatment [163].

Among the children on ART, the rate of viral load monitoring was lower than recommended in the South African guidelines, with individuals tested once every 2 years on average, making the estimation of viral suppression difficult. Testing may have been targeted at those suspected of virological failure, although at least some viral load measurements were found to be missing when data in ACCDB were compared with that recorded in ARTemis. Among those with a measurement available, 70% were suppressed at their most recent test, which is in line with

estimates from other studies of infants in South Africa which range from 56% to 88% [165, 184, 261]. The cumulative incidence of switch to second line by 3 years on ART was 12.4%, and the mortality rate on ART was 2.7 deaths per 100 person-years.

Half of individuals who initiated ART had an interruption to treatment of over 1 month, although the median medication possession ratio (proportion of time with an ART prescription) was 93%. Short, regular treatments breaks may have been caused by drug stock outs, with 9% of clinics in South Africa reporting a stock out of a paediatric drug over a 3 month period in 2015. Alternatively, the requirement for monthly clinic attendance for ART prescriptions may make consistent engagement in care challenging, especially in this low-income setting where there are high level of poverty and unemployment and where individuals may have to travel long distances and pay for public transport to reach healthcare facilities. Multi-monthly prescriptions have been shown to be successful among stable children and adolescents in South Africa, and should be considered here [278].

In Chapter 9, results from throughout the thesis were brought together to estimate the complete cascade of care for infants in the sub-district. By two years of age, 88% of HIV-exposed infants born in the sub-district had been tested, and of the infants diagnosed with HIV, 65% had initiated ART, of whom 53% were still on ART, of whom 68% were virally suppressed. Among the 1,034 children born between June 2010 and December 2016 and still living with HIV in the sub-district on the 30th June 2017, 64% were estimated to have been diagnosed, of whom 30% were estimated to be on ART, of whom 69% were estimated to be virally suppressed. Only one other study, which was conducted prior to the introduction of birth testing, has assessed the complete HIV cascade for infants in South Africa. Of 838 HIV-exposed infants born at a single hospital in Johannesburg, 82% received a PCR test at 6 weeks of age, and of those diagnosed with HIV, 61% started ART, of whom 54% were known to have achieved viral suppression by a median of 18 weeks after treatment initiation [177]. The key limitations for constructing the cascade of care in this chapter were that results were dependent on several stages of linkage, that death and migration prior to ART initiation could only be accounted for among the subset of infants linked to ACDIS, and the low proportion of infants with viral load data. Further, results were dependent on rates of MTCT estimated by mathematical models which varied significantly; estimation of this stage is often the most difficult, with errors here affecting results across the rest of the cascade.

Table 10.1 - Summary of key findings and limitations of each chapter

Chapter and key objective	Key findings	Key limitations
Chapter 4: Data linkage *Link datasets for later use in estimating the cascade of care*	• 19,884 PCR tests extracted from NHLS database, from which 15,234 unique infants (who met the inclusion criteria) were identified • 283 infants with a PCR test linked to ACCDB • 2,349 infants with a PCR test linked to ACDIS	• Mother's name may be recorded instead of infant's • No South African ID or other unique identifier • Limited data items for linkage, especially between NHLS and ACDIS • Limited ways to validate or assess performance of linkage algorithm
Chapter 5: Infant HIV PCR testing coverage *Compare estimates of the proportion of HIV-exposed infants receiving an HIV PCR test from four different methods*	• All the sources of data considered had limitations, which affects the accuracy of their estimates, resulting in high variation between methods • Not all PCR tests conducted were recorded in Road-to-health booklets and DHIS clinic logbooks were likely poorly completed	• NHLS data were missing from one clinic for some of the time period of interest
Chapter 6: Factors associated with HIV PCR testing *Explore associations between key maternal, pregnancy and delivery related characteristics and coverage of HIV PCR testing*	• Conflicting trends over time were observed between the two sources of data used • Other than calendar time, only a longer time since maternal ART initiation and a higher number of previous live-born children associated with increased likelihood of PCR testing	• A high proportion of missing data across some covariates meant they were excluded from analysis • A complete case analysis was used, which, although shown to be unbiased, may have reduced efficiency • Likely misclassification of the outcome would have resulted in attenuation of true effects
Chapter 7: Frequency, timing and results of HIV PCR testing *Estimate the proportion of infants tested in accordance with South African national guidelines and the proportion testing positive*	• Less than a quarter of infants who tested negative at birth received a repeat test • Only half of infants with a positive test result received a confirmatory PCR test • Vaccination coverage was close to 100% for vaccinations in the first year of life, suggesting general engagement in healthcare services was high • Adherence to PCR testing guidelines varied by clinic • Among those tested, the proportion of infants with a positive result fell over time from 6.6% among those born in 2010 to 1.0% in 2016	• Reliant on accurate linkage of multiple tests on the same infants • Vaccination coverage was reported by caregiver, with the majority of infant's Road-to-health booklets not seen, which may result in social desirability bias • Low repeat testing and low overall testing coverage prevented accurate estimation of rate of MTCT

Chapter 8: ART initiation and viral suppression *Estimate the proportion of infants diagnosed with HIV who went on to initiate ART, and key outcomes on ART, including viral suppression*	• Among infants with a positive PCR test result, 53% went on to initiate ART, with those born in more recent years more likely to start • Some infants on ART in ACCDB had no linked positive PCR test; assuming those among this group who were not likely born elsewhere or diagnosed through other means represented missed links between NHLS and ACCDB, an upper bound on ART coverage was calculated to be 77% • Two-thirds of infants initiating treatment were CDC immunological stage 3 • Those diagnosed at birth were less likely to initiate ART than those diagnosed later • Initial regimens not normally recommended for infants were reported • 70% of infants were suppressed at their more recent viral load • Half of infants had an extended interruption to treatment of over one month • By 3 years on ART, the cumulative incidence of switch to second-line was 12.4% • By last follow-up, 26% of those on ART were lost-to-follow-up and 5% were reported to have died	• Estimation of the proportion initiating ART reliant on accurate linkage, with no alternative sources of data available to confirm or assess accuracy of estimate • Quality and completeness of data recorded in TIER.net is unclear, especially initial regimens, laboratory testing and deaths • PCR tests from birth may be more likely to have mother's name recorded making linkage more difficult, leading to underestimation of the proportion of those diagnosed at birth who initiated treatment • Low rate of viral load testing makes interpretation of the proportion achieving viral suppression difficult
Chapter 9: The cascade of care for infants exposed to HIV *Estimate the cascade of care for both HIV-exposed infants and infants with HIV born in the sub-district*	• By two years of age, 88% of HIV-exposed infants born in the sub-district had been tested, and the infants diagnosed with HIV, 65% had initiated ART, of whom 53% were still on ART, of whom 68% were virally suppressed • Among the 1,034 children born between June 2010 and December 2016 and still living with HIV in the sub-district on the 30th June 2017, 64% were estimated to have been diagnosed, of whom 30% were estimated to be on ART, of whom 69% were estimated to be virally suppressed	• Estimates depended on accurate linkage between the three sources of data • No data on migration and data on deaths were only available for the relatively small subset who were linked to demographic surveillance • Low rate of viral load testing makes interpretation of the proportion achieving viral suppression difficult

10.3. Concluding remarks

10.3.1. Strengths and limitations

The specific strengths and limitations of each analysis have been previously described in detail in each chapter, but here I discuss several general strengths and limitations that apply across the thesis as a whole.

Although many studies in South Africa have assessed single stages of the cascade by themselves, only one has considered the complete cascade from HIV acquisition to viral suppression [177]. In addition, this previous study included infants born between 2008 and 2010, before the introduction of birth testing. This thesis is therefore the first time the cascade has been explored for infants in South Africa across a wider population, and over the period where birth testing was policy. The impact of this policy on infant diagnosis and its implications on the later stages of cascade are important to understand, and this thesis therefore addresses some of the current gaps in knowledge about delivering the best care to HIV-exposed infants. Following a single cohort of individuals to assess all cascade stages is preferable to estimation based on different groups of individuals for each stage, or even for the numerator and denominator of a given stage, as it improves the internal consistency, preventing results from being biased by geographical, temporal, or other differences between the groups used.

Few studies using routine health record data consider their limitations to the extent to which has been done for this thesis. This is important both to understand the accuracy of the estimates reported and to identify ways to improve data collections systems. The use of multiple sources of data for the testing stage highlighted the big effect these limitations can have on estimates, which would not be apparent in a study using just a single source.

An important weakness of this research is that although the key most commonly used intermediate cascade stages (diagnosis and ART initiation) were measured, others, such as the return of PCR test results to caregivers and linkage to care prior to ART initiation, were not included as no data were available to assess them. Unfortunately, many of these intermediate stages lie between diagnosis and ART, which is where biggest drop off in cascade retention was seen. Further, the complete cascade for HIV-exposed infants could not be estimated, with no data available on the provision of and adherence to infant ART prophylaxis during breastfeeding or on the receipt of the recommended antibody tests at 18 months of age and the cessation of breastfeeding. Adherence to maternal ART and infant prophylaxis throughout breastfeeding, may be challenging as a result of psychosocial, cultural and economic obstacles [279], with continuous adherence (defined as taking >95% of doses) being estimated at 63% and 75%

respectively to 18 months of age in South Africa [280]; these later tests to detect any new infections are therefore extremely important.

10.3.2. Generalisability

Although South Africa is an upper-middle income country, the Hlabisa health sub-district is a low-income region and so the results here may not be generalisable across the whole of the country, but may be more generalisable to the broader low resource rural regions with the highest prevalence of HIV. Further, only a few high HIV-burden countries, namely Namibia and Botswana, are classified as upper-middle-income, and neither of these recommends routine PCR testing of HIV-exposed infants at birth, although ongoing research is exploring the potential benefits of targeted birth testing [275]. This limits the relevance of the results described here outside of South Africa.

In theory, the methods used to estimate the cascade here could be extended for estimation at a provincial or national level. Although the demographic surveillance data or data from Road-to-health booklets would not be available to estimate the number of HIV-exposed infants, this could be estimated using the data from one of the other methods, and data from NHLS and TIER.net for the later stages are available across South Africa. Limiting factors may be computational difficulties of the number of pairwise comparisons required (which increases to the order of n^2 as the number of infants, n, increases), and the manual review of potential matches in the probabilistic linkage.

10.3.3. Importance of routinely collected data and suggested improvements to data collection

In general, this thesis raises important challenges, but also potential benefits, of good quality routinely collected data and data linkage for the assessment of HIV services. Routinely collected data are convenient and cost-effective compared to collecting data specifically to answer a given research question. They often cover the whole population, and may be longitudinal, allowing the assessment of changes over time. However, limited data items may be available, for example, no data being available in TIER.net on reasons for treatment changes or on linkage to care prior to ART initiation. Further, data items which are collected may not be accurately recorded or may be missing, with no mechanism for data querying, introducing error or selection bias. The WHO promotes the use of case-based surveillance (individual-level tracing of individuals through healthcare services [281]) using data sources such as those used here for both clinical management and for program monitoring, giving recommendations for data quality reviews, including the comparison of data from multiple sources [282].

The findings here lead to several recommendations for improvements to policy and practice, which would improve measurement of the cascade, as well as improve follow-up of individuals in care, both within early infant diagnosis programs as well as other services.

Firstly, the more widespread use of a unique patient identifier should be encouraged. Earlier assignment of the national ID number should be considered, or a separate identification number specific to healthcare services could be used. In the Western Cape, where healthcare is managed by the Western Cape Provincial Department of Health, individuals are assigned a healthcare identifier that is used across laboratory, pharmacy, and service data [283]. As well as enabling access to complete health records through a provincial data centre and thus improving the delivery of clinical care, it has also enabled a breadth of research studies based on linked anonymised data [192, 284]. No comparable identifier is currently available across the rest of the country and while a similar system should be considered, the logistical, financial and technical hurdles involved in its implementation in the other provinces with fewer resources than the Western Cape may result in delays. In the meantime, alternatives could be considered. A pilot study in Tshwane district in Gauteng incorporated a page of pre-printed barcode stickers in newly assigned Road-to-health booklets, which could then be placed on NHLS forms to enable linkage of repeat PCR tests [166]. Limited success was demonstrated in this study, with the barcode recorded on less than half of birth PCR tests, however the idea may warrant further exploration, with improved training of staff in healthcare facilities. AHRI has developed a new data collection system called ClinicLink (which began data collection beyond the time period studied in this thesis, and thus was not used in analysis) to monitor clinic attendance, with individuals visiting clinic asked the reason for their attendance, and data assigned to their ACDIS individual identifier [285]. Refinement of reason categories currently used would however be required to accurately identify infants receiving PCR testing at each visit.

Secondly, reasons for the suspected poor completion of registers and databases, including DHIS clinic logbooks, infant Road-to-health booklets and TIER.net, should be investigated. The importance of the use of these as a tool for monitoring healthcare systems should be emphasised to healthcare workers, and suitable training or modifications to the reporting systems implemented where necessary [286-288]. Thirdly, alternative ways of maintaining accurate and complete records of patient data while mitigating concerns about privacy violations should be explored, for example, investigating the feasibility of using codes in Road-to-health booklets to record HIV-related information [224]. Finally, systems for improved monitoring in the long-term should be explored, including through the rollout of national electronic health records [289, 290]. Additional modules in TIER.net for pre-ART care have been planned since its inception in 2013 [190], although have yet to be widely implemented despite

successful pilot studies (presumably because of capacity concerns), and the possibility of expanding its use to cover services for HIV-exposed infants should also be considered.

10.3.4. Opportunities for further research

There are several important questions that could be investigated in future research. Firstly, across the analyses of PCR testing there was some suggestion that adherence to national guidelines varied by clinic. Time motion studies could be used to identify differences in clinic processes and understand reasons for some clinics having more successful outcomes [291]. Healthcare workers from the better performing clinics could train workers from clinics with less effective practises, leading to overall improvement in healthcare systems within the sub-district.

Secondly, no data were available on several intermediate stages, in particular those in between testing and ART initiation, such as the return of PCR test results to caregivers. Ways to assess these and their impact of poor retention through the cascade should be explored. Given that the Road-to-health booklet is patient-held, recording of results in these implies the return of the result to caregivers; in MONARCH Road-to-health booklet data were only available to the 6 week postnatal visit which may explain the low proportion with a result observed, but a study with a longer duration of follow-up could investigate this further. Alternatively, given that HIV-related stigma may impede the complete recording of data in the booklets, other study designs may have to be considered.

Finally, infants in ACDIS who were exposed to HIV were identified and linked to the NHLS, which allowed an analysis exploring which patient-level factors associated with the receipt of PCR testing. However, given the small number of infants in this group who acquired HIV and then subsequently went on to initiate ART or achieve viral suppression, a comparable analysis of the later stages of the cascade was not possible. Further, limited patient characteristics were available within ACCDB to explore associations here. Research to identify these predictors should be considered, although this may require the collection of data specifically for this purpose.

10.3.5. Emerging issues

There are a number of emerging issues relevant to infants exposed to and living with HIV that may affect both the measurement and the achievement of the cascade of care in the future. Firstly, point-of-care PCR testing has been trialled in several pilot studies in South Africa, and its use is expected to increase in the future [162, 292]. Although point-of-care tests have potential benefits, with fewer opportunities for loss-to-follow-up given the immediate return of results to caregivers and the resulting potential for earlier ART initiation, their use may present challenges for monitoring the first step of the cascade. These tests are not processed through a laboratory,

and so the receipt of testing and the recording of the subsequent results would not be available through NHLS, meaning it may not be possible to estimate testing coverage using these methods from this thesis in the future. Currently, recording would occur in infants Road-to-health booklets, which introduces the potential for multiple errors and biases. A pilot study in the Western Cape successfully digitised results of point-of-care tests conducted at one health centre, which were subsequently linked to the provincial data centre [293]. Similarly, the possibility of the rollout of point-of-care viral load testing in low- and middle-income settings in the future may impact estimation of the last stage of cascade [294].

The scale up of PMTCT processes means that an increasing number of HIV infections now occur in the intrapartum and postpartum period, and thus are not detectable at birth. The proportion of infants who received a repeat test following a negative result at birth was relatively low, suggesting follow-up procedures need improvement, which will become increasingly important. Further, as the rate of MTCT continues to decrease, an increasing number of infants with a positive PCR test result will represent false-positives, and thus the need for high coverage of confirmatory PCR testing will also become more and more important. There is the additional challenge that women are not routinely screened in the postpartum period, so incident HIV infections (and thus newly at risk infants) are not identified. A modelling study, based on HIV epidemics in three countries including South Africa, showed maternal HIV screening at the infant 6 week immunization visit is of good value [295].

10.4. Conclusion

Retention through all stages across the cascade appears low, however limitations of all the sources of data used makes confidence in the accuracy of the results difficult. Despite access to a number sources of data which should have complete coverage and be of high quality, in effect, the cascade of care for infants exposed to HIV cannot currently be estimated. Limitations of data sources such as these are often not well considered in other research. Although the rate of MTCT has fallen over time, primarily as a result of the scale up of maternal testing and treatment, the antenatal seroprevalence in some areas, including the Hlabisa health sub-district, remains close to 50%, resulting in a high number of HIV-exposed infants being born. Mortality among young children with HIV is high, making the achievement of viral suppression and good outcomes on ART extremely important. Although the cascade of care is a useful tool to give an overview of the performance of healthcare systems in achieving this, its use is dependent on good quality data and in general estimates should therefore be interpreted with care. More robust data collection systems are therefore urgently required, both for the quality and continuity of clinical care, as well as for surveillance and research purposes.

References

1. Centers for Disease Control and Prevention (CDC), *Pneumocystis pneumonia--Los Angeles.* MMWR Morbidity and Mortality Weekly Report, 1981. **30**(21): p. 250-2.
2. UNAIDS, *UNAIDS Data 2018*. 2018: Geneva (Switzerland).
3. Cummins, N.W. and Badley, A.D., *Making sense of how HIV kills infected CD4 T cells: implications for HIV cure.* Molecular and Cellular Therapies, 2014. **2**: p. 20-20.
4. Vanhamel, J., Bruggemans, A. and Debyser, Z., *Establishment of latent HIV-1 reservoirs: what do we really know?* Journal of Virus Eradication, 2019. **5**(1): p. 3-9.
5. Kahn, J.O. and Walker, B.D., *Acute human immunodeficiency virus type 1 infection.* New England Journal of Medicine, 1998. **339**(1): p. 33-39.
6. Lewthwaite, P. and Wilkins, E., *Natural history of HIV/AIDS.* Medicine, 2009. **37**(7): p. 333-337.
7. Yarchoan, R., Venzon, D.J., Pluda, J.M., Lietzau, J., Wyvill, K.M., Tsiatis, A.A., et al., *CD4 count and the risk for death in patients infected with HIV receiving antiretroviral therapy.* Annals of Internal Medicine, 1991. **115**(3): p. 184-9.
8. Phillips, A.N., Elford, J., Sabin, C., Bofill, M., Janossy, G. and Lee, C.A., *Immunodeficiency and the risk of death in HIV infection.* JAMA, 1992. **268**(19): p. 2662-6.
9. Sokoya, T., Steel, H.C., Nieuwoudt, M. and Rossouw, T.M., *HIV as a Cause of Immune Activation and Immunosenescence.* Mediators of Inflammation, 2017. **2017**: p. 6825493.
10. Pantaleo, G., Graziosi, C. and Fauci, A.S., *The Immunopathogenesis of Human Immunodeficiency Virus Infection.* New England Journal of Medicine, 1993. **328**(5): p. 327-335.
11. Rodger, A.J., Cambiano, V., Bruun, T., Vernazza, P., Collins, S., van Lunzen, J., et al., *Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy.* JAMA, 2016. **316**(2): p. 171-81.
12. NIH. *FDA Approval of HIV Medicines*. [Access Date: 22nd October 2019]; Available from: https://aidsinfo.nih.gov/understanding-hiv-aids/infographics/25/fda-approval-of-hiv-medicines.
13. Santoro, M.M. and Perno, C.F., *HIV-1 Genetic Variability and Clinical Implications.* ISRN Microbiology, 2013. **2013**: p. 481314-481314.
14. World Health Organization, *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*. 2018, World Health Organization.
15. Insight Start Study Group, *Initiation of antiretroviral therapy in early asymptomatic HIV infection.* New England Journal of Medicine, 2015. **373**(9): p. 795-807.
16. Danel, C., Moh, R., Gabillard, D., Badje, A., Le Carrou, J., Ouassa, T., et al., *A Trial of Early Antiretrovirals and Isoniazid Preventive Therapy in Africa.* New England Journal of Medicine, 2015. **373**(9): p. 808-22.
17. World Health Organization, *Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV*. 2015, World Health Organization.
18. Berger, J., Dunn, J.D., Johnson, M.M., Karst, K.R. and Shear, W.C., *How drug life-cycle management patent strategies may impact formulary management.* American Journal of Managed Care, 2016. **22**(16 Suppl): p. S487-s495.
19. Perez-Casas, C., Mace, C., Berman, D. and Double, J., *Accessing ARVs: untangling the web of price reductions for developing countries*, in *Médecins Sans Frontiers/Campaign for Access to Essential Medicines*. 2001: Geneva.
20. Clinton Health Access Initiative, *ARV market report: the state of the antiretroviral drug market in low-and middle-income countries, 2015–2020*. 2016, Issue.

21. Trickey, A., May, M.T., Vehreschild, J.-J., Obel, N., Gill, M.J., Crane, H.M., et al., *Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: a collaborative analysis of cohort studies.* The Lancet HIV, 2017. **4**(8): p. e349-e356.

22. UNAIDS. *Global AIDS Update 2017 (http://aidsinfo.unaids.org/)*. 2017 Accessed: 28th July 2017].

23. Teasdale, C.A., Marais, B.J. and Abrams, E.J., *HIV: prevention of mother-to-child transmission.* BMJ Clinical Evidence, 2011. **2011**.

24. Kourtis, A.P., Lee, F.K., Abrams, E.J., Jamieson, D.J. and Bulterys, M., *Mother-to-child transmission of HIV-1: timing and implications for prevention.* The Lancet Infectious Diseases, 2006. **6**(11): p. 726-732.

25. Kalish, L.A., Pitt, J., Lew, J., Landesman, S., Diaz, C., Hershow, R., et al., *Defining the time of fetal or perinatal acquisition of human immunodeficiency virus type 1 infection on the basis of age at first positive culture. Women and Infants Transmission Study (WITS).* The Journal of Infectious Diseases, 1997. **175**(3): p. 712-5.

26. Blattner, W., *Women and Infants Transmission Study (WITS) 1990-1999*, in *XIII AIDS Conference*. 2000: Durban, South Africa (LBOr4).

27. Newell, M.L., Dunn, D.T., Peckham, C.S., Semprini, A.E. and Pardi, G., *Vertical transmission of HIV-1: maternal immune status and obstetric factors. The European Collaborative Study.* AIDS (London, England), 1996. **10**(14): p. 1675-81.

28. Chiappini, E., Galli, L., Lisi, C., Gabiano, C., Esposito, S., Giacomet, V., et al., *Strategies for Prevention of Mother-to-Child Transmission Adopted in the "Real-World" Setting: Data From the Italian Register for HIV-1 Infection in Children.* Journal of Acquired Immune Deficiency Syndromes, 2018. **79**(1): p. 54-61.

29. Landesman, S.H., Kalish, L.A., Burns, D.N., Minkoff, H., Fox, H.E., Zorrilla, C., et al., *Obstetrical factors and the transmission of human immunodeficiency virus type 1 from mother to child.* New England Journal of Medicine, 1996. **334**(25): p. 1617-1623.

30. Connor, E.M., Sperling, R.S., Gelber, R., Kiselev, P., Scott, G., O'Sullivan, M.J., et al., *Reduction of Maternal-Infant Transmission of Human Immunodeficiency Virus Type 1 with Zidovudine Treatment.* New England Journal of Medicine, 1994. **331**(18): p. 1173-1180.

31. De Cock, K.M., Fowler, M.G., Mercier, E., de Vincenzi, I., Saba, J., Hoff, E., et al., *Prevention of Mother-to-Child HIV Transmission in Resource-Poor CountriesTranslating Research Into Policy and Practice.* JAMA, 2000. **283**(9): p. 1175-1182.

32. World Health Organization, *Antiretroviral Drugs for Treating Pregnant Women and Preventing HIV Infection in Infants. Recommendations for a Public Health Approach*. 2010: Geneva, Switzerland.

33. World Health Organization, *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infections: Recommendations for a Public Health Approach*. 2013: Geneva, Switzerland.

34. World Health Organization, *Use of Antiretroviral Drugs for Treating Pregnant Women and Preventing HIV Infection in Infants*. 2012: Geneva, Switzerland.

35. Victoria, C., *Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis.* The Lancet, 2000. **355**(9202): p. 451-455.

36. Coutsoudis, A., Pillay, K., Kuhn, L., Spooner, E., Tsai, W.Y. and Coovadia, H.M., *Method of feeding and transmission of HIV-1 from mothers to children by 15 months of age: prospective cohort study from Durban, South Africa.* AIDS (London, England), 2001. **15**(3): p. 379-87.

37. Kuhn, L. and Stein, Z.A., *Mother-to-infant HIV transmission: timing, risk factors and prevention.* Paediatric and Perinatal Epidemiology, 1995. **9**(1): p. 1-29.

38. Peters, H., Francis, K., Sconza, R., Horn, A., S. Peckham, C., Tookey, P.A., et al., *UK Mother-to-Child HIV Transmission Rates Continue to Decline: 2012–2014.* Clinical Infectious Diseases, 2017. **64**(4): p. 527-528.

39. UNAIDS and PEPFAR, *On the Fast-Track to an AIDS-free generation*. 2016: Geneva, Switzerland.

40. WHO., *Global guidance on criteria and processes for validation: Elimination of Mother-to-Child Transmission of HIV and Syphilis*. 2014, WHO.: Geneva, Switzerland.

41. World Health Organization. *WHO validation for the elimination of mother-to-child transmission of HIV and/or syphilis*. Available from: https://www.who.int/reproductivehealth/congenital-syphilis/WHO-validation-EMTCT/en/ (last accessed 29/11/19).

42. The European Collaborative Study, *Mother-to-child transmission of HIV infection.* The Lancet, 1988. **2**(8619): p. 1039-43.

43. WHO, *Consolidated Guidelines on the Use of Antiretroviral Drugs for the Treating and Preventing HIV Infection: recommendations for a public health approach*. 2016, World Health Organisation 2016: Geneva.

44. Penazzato, M., Revill, P., Prendergast, A.J., Collins, I.J., Walker, S., Elyanu, P.J., et al., *Early infant diagnosis of HIV infection in low-income and middle-income countries: does one size fit all?* The Lancet Infectious Diseases, 2014. **14**(7): p. 650-655.

45. Rouzioux, C. and Avettand-Fenoël, V., *Total HIV DNA: a global marker of HIV persistence.* Retrovirology, 2018. **15**(1): p. 30-30.

46. Burgard, M., Blanche, S., Jasseron, C., Descamps, P., Allemon, M.-C., Ciraru-Vigneron, N., et al., *Performance of HIV-1 DNA or HIV-1 RNA Tests for Early Diagnosis of Perinatal HIV-1 Infection during Anti-Retroviral Prophylaxis.* The Journal of Pediatrics, 2012. **160**(1): p. 60-66.e1.

47. Connolly, M.D., Rutstein, R.M. and Lowenthal, E.D., *Virologic testing in infants with perinatal exposure to HIV receiving multidrug prophylaxis.* The Pediatric Infectious Disease Journal, 2013. **32**(2): p. 196-197.

48. Sherman, G.G., Cooper, P.A., Coovadia, A.H., Puren, A.J., Jones, S.A., Mokhachane, M., et al., *Polymerase Chain Reaction for Diagnosis of Human Immunodeficiency Virus Infection in Infancy in Low Resource Settings.* The Pediatric Infectious Disease Journal, 2005. **24**(11): p. 993-997.

49. Lilian, R.R., Kalk, E., Bhowan, K., Berrie, L., Carmona, S., Technau, K., et al., *Early diagnosis of in utero and intrapartum HIV infection in infants prior to 6 weeks of age.* Journal of Clinical Microbiology, 2012. **50**(7): p. 2373-7.

50. Hsiao, N.-y., Dunning, L., Kroon, M. and Myer, L., *Laboratory Evaluation of the Alere q Point-of-Care System for Early Infant HIV Diagnosis.* PLOS One, 2016. **11**(3): p. e0152672.

51. Diallo, K., Modi, S., Hurlston, M., Beard, R.S. and Nkengasong, J.N., *A proposed framework for the implementation of early infant diagnosis point-of-care.* AIDS Research and Human Retroviruses, 2017. **33**(3): p. 203-210.

52. Jani, I.V., Meggi, B., Loquiha, O., Tobaiwa, O., Mudenyanga, C., Zitha, A., et al., *Effect of point-of-care early infant diagnosis on antiretroviral therapy initiation and retention of patients.* AIDS (London, England), 2018. **32**(11): p. 1453-1463.

53. Mwenda, R., Fong, Y., Magombo, T., Saka, E., Midiani, D., Mwase, C., et al., *Significant Patient Impact Observed Upon Implementation of Point-of-Care Early Infant Diagnosis Technologies in an Observational Study in Malawi.* Clinical Infectious Diseases, 2018. **67**(5): p. 701-707.

54. Dube, Q., Dow, A., Chirambo, C., Lebov, J., Tenthani, L., Moore, M., et al., *Implementing early infant diagnosis of HIV infection at the primary care level: experiences and challenges in Malawi.* Bulletin of the World Health Organization, 2012. **90**(9): p. 699-704.

55. Violari, A., Cotton, M.F., Gibb, D.M., Babiker, A.G., Steyn, J., Madhi, S.A., et al., *Early antiretroviral therapy and mortality among HIV-infected infants.* New England Journal of Medicine, 2008. **359**(21): p. 2233-2244.

56. Persaud, D., Gay, H., Ziemniak, C., Chen, Y.H., Piatak, M., Jr., Chun, T.W., et al., *Absence of detectable HIV-1 viremia after treatment cessation in an infant.* New England Journal of Medicine, 2013. **369**(19): p. 1828-35.

57. Lewis, J., Walker, A.S., Castro, H., De Rossi, A., Gibb, D.M., Giaquinto, C., et al., *Age and CD4 count at initiation of antiretroviral therapy in HIV-infected children: effects on long-term T-cell reconstitution.* Journal of Infectious Diseases, 2012. **205**(4): p. 548-56.

58. Laughton, B., Cornell, M., Kidd, M., Springer, P.E., Dobbels, E.F.M., Rensburg, A.J.V., et al., *Five year neurodevelopment outcomes of perinatally HIV-infected children on early limited or deferred continuous antiretroviral therapy.* Journal of the International AIDS Society, 2018. **21**(5): p. e25106.

59. McGrath, C.J., Diener, L., Richardson, B.A., Peacock-Chambers, E. and John-Stewart, G.C., *Growth reconstitution following antiretroviral therapy and nutritional supplementation: systematic review and meta-analysis.* AIDS (London, England), 2015. **29**(15): p. 2009-2023.

60. Paediatric European Network for Treatment of AIDS (PENTA), *Once vs. twice-daily lopinavir/ritonavir in HIV-1-infected children.* AIDS, 2015. **29**(18): p. 2447-57.

61. Paediatric European Network for Treatment of AIDS (PENTA), *Response to planned treatment interruptions in HIV infection varies across childhood.* AIDS, 2010. **24**(2): p. 231-241.

62. Cotton, M.F., Violari, A., Otwombe, K., Panchia, R., Dobbels, E., Rabie, H., et al., *Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the children with HIV early antiretroviral (CHER) randomised trial.* The Lancet, 2013. **382**(9904): p. 1555-1563.

63. The BREATHER (PENTA 16) Trial Group, *Weekends-off efavirenz-based antiretroviral therapy in HIV-infected children, adolescents, and young adults (BREATHER): a randomised, open-label, non-inferiority, phase 2/3 trial.* The Lancet HIV, 2016. **3**(9): p. e421-e430.

64. Penazzato, M., Palladino, C. and Sugandhi, N., *Prioritizing the most needed formulations to accelerate paediatric antiretroviral therapy scale-up.* Current Opinion in HIV and AIDS, 2017. **12**(4): p. 369-376.

65. Best, B.M., Capparelli, E.V., Diep, H., Rossi, S.S., Farrell, M.J., Williams, E., et al., *Pharmacokinetics of lopinavir/ritonavir crushed versus whole tablets in children.* Journal of Acquired Immune Deficiency Syndromes, 2011. **58**(4): p. 385-91.

66. Penazzato, M., Gnanashanmugam, D., Rojo, P., Lallemant, M., Lewis, L.L., Rocchi, F., et al., *Optimizing Research to Speed Up Availability of Pediatric Antiretroviral Drugs and Formulations.* Clinical Infectious Diseases, 2017. **64**(11): p. 1597-1603.

67. Crichton, S., Collins, I.J., Bamford, A., Doerholt, K., Riordan, A., Lyall, H., et al., *Abacavir use in young infants in the UK and Ireland national paediatric cohort*, in *11th Workshop on HIV Pediatrics*. 2019: Mexico City, Mexico.

68. Bamford, A., Turkova, A., Lyall, H., Foster, C., Klein, N., Bastiaans, D., et al., *Paediatric European Network for Treatment of AIDS (PENTA) guidelines for treatment of paediatric HIV-1 infection 2015: optimizing health in preparation for adult life.* HIV Medicine, 2018. **19**(1): p. e1-e42.

69. Tao, X., Lu, Y., Zhou, Y., Huang, Y. and Chen, Y., *Virologically suppressed HIV-infected patients on TDF-containing regimens significantly benefit from switching to TAF-containing regimens: A meta-analysis of randomized controlled trials.* International Journal of Infectious Diseases, 2019. **87**: p. 43-53.

70. Paediatric European Network for Treatment of AIDS (PENTA), *Comparison of dual nucleoside-analogue reverse-transcriptase inhibitor regimens with and without nelfinavir in children with HIV-1 who have not previously been treated: the PENTA 5 randomised trial.* The Lancet, 2002. **359**(9308): p. 733-740.

71. PENPACT-1 Study Team, Babiker, A., Castro nee Green, H., Compagnucci, A., Fiscus, S., Giaquinto, C., et al., *First-line antiretroviral therapy with a protease inhibitor versus non-nucleoside reverse transcriptase inhibitor and switch at higher versus low viral load in HIV-infected children: an open-label, randomised phase 2/3 trial.* The Lancet Infectious Diseases, 2011. **11**(4): p. 273-283.

72. Barlow-Mosha, L., Angelidou, K., Lindsey, J., Archary, M., Cotton, M., Dittmer, S., et al., *Nevirapine- Versus Lopinavir/Ritonavir-Based Antiretroviral Therapy in HIV-Infected Infants and Young Children: Long-term Follow-up of the IMPAACT P1060 Randomized Trial.* Clinical Infectious Diseases, 2016. **63**(8): p. 1113-1121.

73. McArthur, M.A., Kalu, S.U., Foulks, A.R., Aly, A.M., Jain, S.K. and Patel, J.A., *Twin preterm neonates with cardiac toxicity related to lopinavir/ritonavir therapy.* The Pediatric Infectious Disease Journal, 2009. **28**(12): p. 1127-1129.

74. Simon, A., Warszawski, J., Kariyawasam, D., Le Chenadec, J., Benhammou, V., Czernichow, P., et al., *Association of prenatal and postnatal exposure to lopinavir-ritonavir and adrenal dysfunction among uninfected infants of HIV-infected mothers.* JAMA, 2011. **306**(1): p. 70-78.

75. Prendergast, A.J., Klenerman, P. and Goulder, P.J., *The impact of differential antiviral immunity in children and adults.* Nature Reviews Immunology, 2012. **12**(9): p. 636-48.

76. Ásbjörnsdóttir, K.H., Hughes, J.P., Wamalwa, D., Langat, A., Slyker, J.A., Okinyi, H.M., et al., *Differences in virologic and immunologic response to antiretroviral therapy among HIV-1-infected infants and children.* AIDS (London, England), 2016. **30**(18): p. 2835-2843.

77. Judd, A. and on behalf of the European Pregnancy Paediatric HIV Cohort Collaboration study group in EuroCoord, *Early antiretroviral therapy in HIV-1-infected infants, 1996-2008: treatment response and duration of first-line regimens.* AIDS (London, England), 2011. **25**(18): p. 2279-2287.

78. Jenabian, M.-A., Costiniuk, C.T., Mboumba Bouassa, R.-S., Chapdeleine Mekue Mouafo, L., Brogan, T.V. and Bélec, L., *Tackling virological failure in HIV-infected children living in Africa.* Expert Review of Anti-infective Therapy, 2015. **13**(10): p. 1213-1223.

79. The European Collaborative Study, *Age-related standards for T lymphocyte subsets based on uninfected children born to human immunodeficiency virus 1-infected women*.* The Pediatric Infectious Disease Journal, 1992. **11**(12): p. 1018-1026.

80. Centers for Disease Control and Prevention, *Revised surveillance case definition for HIV infection--United States, 2014.* MMWR Morbidity and Mortality Weekly Report, 2014. **63**(Rr-03): p. 1-10.

81. Meyers, T., Moultrie, H., Naidoo, K., Cotton, M., Eley, B. and Sherman, G., *Challenges to pediatric HIV care and treatment in South Africa.* The Journal of Infectious Diseases, 2007. **196**(Supplement_3): p. S474-S481.

82. Marinda, E., Humphrey, J.H., Iliff, P.J., Mutasa, K., Nathoo, K.J., Piwoz, E.G., et al., *Child mortality according to maternal and infant HIV status in Zimbabwe.* The Pediatric Infectious Disease Journal, 2007. **26**(6): p. 519-26.

83. Abrams, E.J., Wiener, J., Carter, R., Kuhn, L., Palumbo, P., Nesheim, S., et al., *Maternal health factors and early pediatric antiretroviral therapy influence the rate of perinatal HIV-1 disease progression in children.* AIDS, 2003. **17**(6): p. 867-877.

84. Ben-Farhat, J., Schramm, B., Nicolay, N., Wanjala, S., Szumilin, E., Balkan, S., et al., *Mortality and clinical outcomes in children treated with antiretroviral therapy in four African vertical programmes during the first decade of paediatric HIV care, 2001–2010.* Tropical Medicine & International Health, 2017. **22**(3): p. 340-350.

85. The European Pregnancy Paediatric HIV Cohort Collaboration study group in EuroCoord, Judd, A., Chappell, E., Turkova, A., Le Coeur, S., Noguera-Julian, A., et al., *Long-term trends in mortality and AIDS-defining events after combination ART initiation among children and adolescents with perinatal HIV infection in 17 middle- and high-income countries in Europe and Thailand: A cohort study.* PLOS Medicine, 2018. **15**(1): p. e1002491.

86. World Health Organization, *WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children*. 2007, World Health Organization: Geneva.

87. Tanser, F., Barnighausen, T., Grapsa, E., Zaidi, J. and Newell, M.L., *High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa.* Science, 2013. **339**(6122): p. 966-71.

88. Wilber, J.A. and Barrow, J.G., *Hypertension—A community problem.* The American Journal of Medicine, 1972. **52**(5): p. 653-663.

89. Gardner, E.M., McLees, M.P., Steiner, J.F., Del Rio, C. and Burman, W.J., *The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection.* Clinical Infectious Diseases, 2011. **52**(6): p. 793-800.

90. UNAIDS, *90-90-90: an ambitious treatment target to help end the AIDS epidemic*. 2014: Geneva (Switzerland).

91. Gisslén, M., Svedhem, V., Lindborg, L., Flamholc, L., Norrgren, H., Wendahl, S., et al., *Sweden, the first country to achieve the Joint United Nations Programme on HIV/AIDS (UNAIDS)/World Health Organization (WHO) 90-90-90 continuum of HIV care targets.* HIV Medicine, 2017. **18**(4): p. 305-307.

92. Haber, N., Pillay, D., Porter, K. and Barnighausen, T., *Constructing the cascade of HIV care: methods for measurement.* Current Opinion in HIV and AIDS, 2016. **11**(1): p. 102-8.

93. Gourlay, A.J., Pharris, A.M., Noori, T., Supervie, V., Rosinska, M., van Sighem, A., et al., *Towards standardised definitions for monitoring the continuum of HIV care in Europe.* AIDS, 2017. **31**(15): p. 2053-2058.

94. Lourenco, L., Hull, M., Nosyk, B., Montaner, J.S. and Lima, V.D., *The need for standardisation of the HIV continuum of care.* The Lancet HIV, 2015. **2**(6): p. e225-6.

95. International Advisory Panel on HIV Care Continuum Optimization, *IAPAC Guidelines for Optimizing the HIV Care Continuum for Adults and Adolescents.* Journal of the International Association of Providers of AIDS Care, 2015. **14 Suppl 1**: p. S3-S34.

96. WHO, *Consolidated Strategic Information Guidelines for HIV in the Health Sector*. 2015, World Health Organization 2015.: Geneva.

97. Granich, R., Gupta, S., Hall, I., Aberle-Grasse, J., Hader, S. and Mermin, J., *Status and methodology of publicly available national HIV care continua and 90-90-90 targets: A systematic review.* PLOS Medicine, 2017. **14**(4): p. e1002253.

98. Maman, D., Zeh, C., Mukui, I., Kirubi, B., Masson, S., Opolo, V., et al., *Cascade of HIV care and population viral suppression in a high-burden region of Kenya.* AIDS (London, England), 2015. **29**(12): p. 1557.

99. Nosyk, B., Montaner, J.S.G., Colley, G., Lima, V.D., Chan, K., Heath, K., et al., *The cascade of HIV care in British Columbia, Canada, 1996–2011: a population-based retrospective cohort study.* The Lancet Infectious Diseases, 2014. **14**(1): p. 40-49.

100. Alvarez-Uria, G., Pakam, R., Midde, M. and Naik, P.K., *Entry, Retention, and Virological Suppression in an HIV Cohort Study in India: Description of the Cascade of Care and Implications for Reducing HIV-Related Mortality in Low- and Middle-Income Countries.* Interdisciplinary Perspectives on Infectious Diseases, 2013. **2013**: p. 384805.

101. Chappell, E., Lyall, H., Riordan, A., Thorne, C., Foster, C., Butler, K., et al., *The cascade of care for children and adolescents with HIV in the UK and Ireland, 2010 to 2016.* Journal of the International AIDS Society, 2019. **22**(9): p. e25379.

102. Haber, N., Tanser, F., Bor, J., Naidu, K., Mutevedzi, T., Herbst, K., et al., *From HIV infection to therapeutic response: a population-based longitudinal HIV cascade-of-care study in KwaZulu-Natal, South Africa.* The Lancet HIV, 2017. **4**(5): p. e223-e230.

103. Lazarus, J.V., Safreed-Harmon, K., Barton, S.E., Costagliola, D., Dedes, N., del Amo Valero, J., et al., *Beyond viral suppression of HIV–the new quality of life frontier.* BMC Medicine, 2016. **14**(1): p. 94.

104. Bärnighausen, T., *The HIV treatment cascade and antiretroviral impact in different populations.* Current Opinion in HIV and AIDS, 2015. **10**(6): p. 391-394.

105. Statistics South Africa, *Census 2011: Census in brief*. 2011, Statistics South Africa: Pretoria. p. p18.

106. The World Bank, *Overcoming Poverty and Inequality in South Africa: An Assessment of Drivers, Constraints and Opportunities*. 2018: Washington DC, USA.

107. Ataguba, J.E., Akazili, J. and McIntyre, D., *Socioeconomic-related health inequality in South Africa: evidence from General Household Surveys.* International Journal for Equity in Health, 2011. **10**(1): p. 48.

108. Coovadia, H., Jewkes, R., Barron, P., Sanders, D. and McIntyre, D., *The health and health system of South Africa: historical roots of current public health challenges.* The Lancet, 2009. **374**(9692): p. 817-834.

109. United Nations Department of Economic and Social Affairs Population Division, *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables*. 2015: New York, USA.

110. National Department of Health (NDoH), Statistics South Africa (Stats SA), South African Medical Research Council (SAMRC) and ICF, *South Africa Demographic and Health Survey 2016: Key Indicators*. 2017: Pretoria, South Africa and Rockville, Maryland, USA.

111. Simelela, N. and Venter, W.D.F., *A brief history of South Africa's response to AIDS.* South African Medical Journal, 2014. **104**(3): p. 249-251.

112. Nattrass, N., *AIDS and the Scientific Governance of Medicine in Post-Apartheid South Africa.* African Affairs, 2008. **107**(427): p. 157-176.

113. Johnson, L.F., May, M.T., Dorrington, R.E., Cornell, M., Boulle, A., Egger, M., et al., *Estimating the impact of antiretroviral treatment on adult mortality trends in South Africa: A mathematical modelling study.* PLOS Medicine, 2017. **14**(12): p. e1002468.

114. South African National AIDS Council, *South Africa Global AIDS Response Progress Report (GARPR)*. 2015: Johannesburg.

115. Zaidi, J., Grapsa, E., Tanser, F., Newell, M.L. and Barnighausen, T., *Dramatic increase in HIV prevalence after scale-up of antiretroviral treatment.* AIDS, 2013. **27**(14): p. 2301-5.

116. South African National Department of Health, *Uniform Patient Fee Schedule (UPFS) for Paying Patients Attending Public Hospitals: User Guide*. 2015.

117. Burton, R., Giddy, J. and Stinson, K., *Prevention of mother-to-child transmission in South Africa: an ever-changing landscape.* Obstetric Medicine, 2015. **8**(1): p. 5-12.

118. National Department of Health of South Africa, *Policy and Guidelines for the Implementation of the PMTCT programme*. 2008: National Department of Health, Pretoria.

119. National Department of Health of South Africa, *Clinical Guidelines: PMTCT (Prevention of Mother to Child Transmission)*. 2010: National Department of Health, Pretoria.

120. National Department of Health of South Africa, *The South African Antiretroviral Treatment Guidelines*. 2013: National Department of Health, Pretoria.

121. National Department of Health of South Africa, *National Consolidated Guidelines For The Prevention Of Mother-To-Child Transmission Of HIV (PMTCT) And The Management Of HIV In Children, Adolescents And Adults.* 2015: National Department of Health, Pretoria.

122. Moyo, F., Mazanderani, A.H., Barron, P., Bhardwaj, S., Goga, A.E., Pillay, Y., et al., *Introduction of Routine HIV Birth Testing in the South African National Consolidated Guidelines.* The Pediatric Infectious Disease Journal, 2018. **37**(6): p. 559-563.

123. Sherman, G.G., Mazanderani, A.H., Barron, P., Bhardwaj, S., Niit, R., Okobi, M., et al., *Toward elimination of mother-to-child transmission of HIV in South Africa: how best to monitor early infant infections within the Prevention of Mother-to-Child Transmission Program.* Journal of Global Health, 2017. **7**(1): p. 010701.

124. Coetzee, D., Hildebrand, K., Boulle, A., Maartens, G., Louis, F., Labatala, V., et al., *Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa.* AIDS, 2004. **18**(6): p. 887-95.

125. National Department of Health of South Africa, *National Antiretroviral Treatment Guidelines*. 2004: National Department of Health, Pretoria.

126. National Department of Health of South Africa, *The South African Antiretroviral Treatment Guidelines*. 2010: National Department of Health, Pretoria.

127. UNAIDS. *South Africa takes bold step to provide HIV treatment for all (http://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2016/may/20160513_UTT)*. 2016 Accessed: 28th July 2017].

128. National Department of Health of South Africa, *Vaccinator's Manual*. 2010: Pretoria.

129. Saloojee, H., Gray, G. and McIntyre, J., *HIV and infant feeding-one step forwards, two steps back: opinion.* Southern African Journal of HIV Medicine, 2011. **12**(4): p. 6-10.

130. Larmarange, J., Mossong, J., Barnighausen, T. and Newell, M.L., *Participation dynamics in population-based longitudinal HIV surveillance in rural South Africa.* PLOS One, 2015. **10**(4): p. e0123345.

131. McGrath, N., Eaton, J.W., Newell, M.-L. and Hosegood, V., *Migration, sexual behaviour, and HIV risk: a general population cohort in rural South Africa.* The Lancet HIV, 2015. **2**(6): p. e252-e259.

132. Houlihan, C.F., Bland, R.M., Mutevedzi, P.C., Lessells, R.J., Ndirangu, J., Thulare, H., et al., *Cohort profile: Hlabisa HIV treatment and care programme.* International Journal of Epidemiology, 2011. **40**(2): p. 318-26.

133. Tanser, F., Hosegood, V., Bärnighausen, T., Herbst, K., Nyirenda, M., Muhwava, W., et al., *Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey.* International Journal of Epidemiology, 2008. **37**(5): p. 956-962.

134. *Central Chronic Medicine Dispensing and Distribution Programme*. [cited Last accessed: 10th February 2018]; Available from: https://za.usembassy.gov/wp-content/uploads/sites/19/2016/06/Central-Chronic-Medicine-Dispensing-and-Distribution-Programme_Ricardo-Kettledas.pdf.

135. Tanser, F., Gijsbertsen, B. and Herbst, K., *Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system.* Social Science & Medicine, 2006. **63**(3): p. 691-705.

136. Vandormael, A., de Oliveira, T., Tanser, F., Bärnighausen, T. and Herbeck, J.T., *High percentage of undiagnosed HIV cases within a hyperendemic South African community: a population-based study.* Journal of Epidemiology and Community Health, 2018. **72**(2): p. 168-172.

137. Vandormael, A., Akullian, A., Siedner, M., de Oliveira, T., Bärnighausen, T. and Tanser, F., *Declines in HIV incidence among men and women in a South African population-based cohort.* Nature Communications, 2019. **10**(1): p. 5482.

138. Yapa, H., de Neve, J., Chetty, T., Herbst, C., Post, F., Cooper, D., et al. *Does Continuous Quality Improvement Improve PMTCT Processes in rural South Africa? A Stepped-Wedge Cluster RCT*. in *CROI 2018*. 2018. Boston, USA.

139. Reniers, G., Blom, S., Calvert, C., Martin-Onraet, A., Herbst, A.J., Eaton, J.W., et al., *Trends in the burden of HIV mortality after roll-out of antiretroviral therapy in KwaZulu-Natal, South Africa: an observational community cohort study.* The Lancet HIV, 2017. **4**(3): p. e113-e121.

140. Tlou, B., Sartorius, B. and Tanser, F., *Space-time variations in child mortality in a rural South African population with high HIV prevalence (2000–2014).* PLOS One, 2017. **12**(8): p. e0182478.

141. Tlou, B., Sartorius, B. and Tanser, F., *Investigating risk factors for under-five mortality in an HIV hyper-endemic area of rural South Africa, from 2000-2014.* PLOS ONE, 2018. **13**(11): p. e0207294.

142. Price, J., Willcox, M., Kabudula, C.W., Herbst, K., Kahn, K. and Harnden, A., *Home deaths of children under 5 years in rural South Africa: a population-based longitudinal study.* Tropical Medicine & International Health, 2019. **24**(7): p. 862-878.

143. Janssen, N., Ndirangu, J., Newell, M.-L. and Bland, R.M., *Successful paediatric HIV treatment in rural primary care in Africa.* Archives of Disease in Childhood, 2009: p. adc. 2009.169367.

144. Cooke, G.S., Little, K.E., Bland, R.M., Thulare, H. and Newell, M.-L., *Need for Timely Paediatric HIV Treatment within Primary Health Care in Rural South Africa.* PLOS One, 2009. **4**(9): p. e7101.

145. Bland, R.M., Ndirangu, J. and Newell, M.-L., *Maximising opportunities for increased antiretroviral treatment in children in an existing HIV programme in rural South Africa.* British Medical Journal, 2013. **346**: p. f550.

146. Pillay, S., Bland, R.M., Lessells, R.J., Manasa, J., de Oliveira, T. and Danaviah, S., *Drug resistance in children at virological failure in a rural KwaZulu-Natal, South Africa, cohort.* AIDS Research and Therapy, 2014. **11**(1): p. 3.

147. Chetty, T., Thorne, C., Tanser, F., Barnighausen, T. and Coutsoudis, A., *Cohort profile: the Hlabisa pregnancy cohort, KwaZulu-Natal, South Africa.* BMJ Open, 2016. **6**(10): p. e012088.

148. Chetty, T., Vandormael, A., Thorne, C. and Coutsoudis, A., *Incident HIV during pregnancy and early postpartum period: a population-based cohort study in a rural area in KwaZulu-Natal, South Africa.* BMC Pregnancy Childbirth, 2017. **17**(1): p. 248.

149. Bland, R., Coovadia, H., Coutsoudis, A., Rollins, N. and Newell, M., *Cohort profile: mamanengane or the Africa centre vertical transmission study.* International Journal of Epidemiology, 2010. **39**(2): p. 351-60.

150. Coovadia, H.M., Rollins, N.C., Bland, R.M., Little, K., Coutsoudis, A., Bennish, M.L., et al., *Mother-to-child transmission of HIV-1 infection during exclusive breastfeeding in the first 6 months of life: an intervention cohort study.* The Lancet, 2007. **369**(9567): p. 1107-1116.

151. Chappell, E., Baisley, K., Bärnighausen, T., Collins, I.J., Gareta, D., Gibb, D., et al., *Impact of the introduction of HIV testing at birth on early infant diagnosis in KwaZulu-Natal, South Africa, 2010-2017*, in *10th HIV Paediatrics Workshop*. 2018: Amsterdam, Netherlands.

152. Chappell, E., Baisley, K., Bärnighausen, T., Collins, I.J., Gareta, D., Herbst, K., et al., *Viral suppression, ART interruptions and switching among children with HIV on ART in KwaZulu-Natal, South Africa, 2010-2016*, in *11th HIV Paediatrics Workshop*. 2019: Mexico City, Mexico.

153. Geddes, R., Knight, S., Reid, S., Giddy, J., Esterhuizen, T. and Roberts, C., *Prevention of mother-to-child transmission of HIV programme: low vertical transmission in KwaZulu-Natal, South Africa.* South African Medical Journal, 2008. **98**(6): p. 458-62.

154. Feucht, U.D., Meyer, A., Thomas, W.N., Forsyth, B.W. and Kruger, M., *Early diagnosis is critical to ensure good outcomes in HIV-infected children: outlining barriers to care.* AIDS Care, 2016. **28**(1): p. 32-42.

155. Technau, K.G., Kuhn, L., Coovadia, A., Murnane, P.M. and Sherman, G., *Xpert HIV-1 point-of-care test for neonatal diagnosis of HIV in the birth testing programme of a maternity hospital: a field evaluation study.* The Lancet HIV, 2017. **4**(10): p. e442-e448.

156. Doherty, T.M., McCoy, D. and Donohue, S., *Health system constraints to optimal coverage of the prevention of mother-to-child HIV transmission programme in South Africa: lessons from the implementation of the national pilot programme.* African Health Sciences, 2005. **5**(3): p. 213-8.

157. Smith, S., Govender, K., Moodley, P., La Russa, P., Kuhn, L. and Archary, M., *Impact of Shifts to Birth Testing on Early Infant Diagnosis Program Outcomes in KwaZulu-Natal, South Africa.* The Pediatric Infectious Disease Journal, 2019. **38**(7): p. e138-e142.

158. Smith, S., Govender, K., Moodley, P., La Russa, P., Kuhn, L. and M., A. *Impact of Shifts to Birth Testing on Early Infant Diagnosis Program Outcomes in KwaZulu-Natal, South Africa*. in *CROI 2018*. 2018. Boston, USA.

159. Moodley, P., Parboosing, R. and Moodley, D., *Reduction in perinatal HIV infections in KwaZulu-Natal, South Africa, in the era of more effective prevention of mother to child transmission interventions (2004-2012).* Journal of Acquired Immune Deficiency Syndromes (1999), 2013. **63**(3): p. 410-5.

160. Kalk, E., Kroon, M., Boulle, A., Osler, M., Euvrard, J., Stinson, K., et al., *Neonatal and infant diagnostic HIV-PCR uptake and associations during three sequential policy periods in Cape Town, South Africa: a longitudinal analysis.* Journal of the International AIDS Society, 2018. **21**(11): p. e25212.

161. Maritz, J., Hsiao, N.-y., Preiser, W. and Myer, L. *Low uptake of routine infant diagnostic testing following HIV PCR testing at birth*. in *Conference on Retroviruses and Opportunistic Infections, Boston, Massachusetts*. 2016.

162. Moyo, F., Chakezha, T., Haeri Mazanderani, A., Sherman, G., Murray, T., Mukendi, A., et al., *Outcomes of point-of-care maternal HIV viral load testing and early infant diagnosis at delivery: experience from four tertiary obstetric units in Gauteng, South Africa*, in *11th International Workshop on HIV Pediatrics*. 2019: Mexico City, Mexico.

163. Technau, K.G., Kuhn, L., Coovadia, A., Carmona, S. and Sherman, G., *Improving early identification of HIV-infected neonates with birth PCR testing in a large urban hospital in Johannesburg, South Africa: successes and challenges.* Journal of the International AIDS Society, 2017. **20**(1): p. 1-8.

164. Nelson, A., Trivino Duran, L., Cassidy, T., Van Cutsem, G., Steele, S.J., Cotton, M., et al., *Impact of Birth PCR on Retention in Care of HIV-Exposed Infants in Primary Care*, in *CROI 2017*. 2017: Seattle, USA.

165. Dunning, L., Kroon, M., Fourie, L., Ciaranello, A. and Myer, L., *Impact of Birth HIV-PCR Testing on the Uptake of Follow-up Early Infant Diagnosis Services in Cape Town, South Africa.* The Pediatric Infectious Disease Journal, 2017. **36**(12): p. 1159-1164.

166. Haeri Mazanderani, A., Sherman, G.G., Moyo, F., Goga, A.E. and Feucht, U., *Leveraging the Road to Health booklet as a unique patient identifier to monitor the prevention of mother-to-child transmission programme.* South African Medical Journal, 2018. **108**(9): p. 729-733.

167. Diallo, K., Kim, A.A., Lecher, S., Ellenberger, D., Beard, R.S., Dale, H., et al., *Early Diagnosis of HIV Infection in Infants - One Caribbean and Six Sub-Saharan African Countries, 2011-2015.* CDC MMWR Morbidity and Mortality Weekly Report, 2016. **65**(46): p. 1285-1290.

168. Sherman, G.G., Lilian, R.R., Bhardwaj, S., Candy, S. and Barron, P., *Laboratory information system data demonstrate successful implementation of the prevention of mother-to-child transmission programme in South Africa.* South African Medical Journal, 2014. **104**(3 Suppl 1): p. 235-8.

169. Feucht, U.D., Meyer, A. and Kruger, M., *Missing HIV prevention opportunities in South African children--a 7-year review.* BMC Public Health, 2014. **14**: p. 1265.

170. Horwood, C., Haskins, L., Vermaak, K., Phakathi, S., Subbaye, R. and Doherty, T., *Prevention of mother to child transmission of HIV (PMTCT) programme in KwaZulu-Natal, South Africa: an evaluation of PMTCT implementation and integration into routine maternal, child and women's health services.* Tropical Medicine & International Health, 2010. **15**(9): p. 992-9.

171. Schwartz, S.R., Clouse, K., Yende, N., Van Rie, A., Bassett, J., Ratshefola, M., et al., *Acceptability and Feasibility of a Mobile Phone-Based Case Management Intervention to Retain Mothers and Infants from an Option B+ Program in Postpartum HIV Care.* Maternal and Child Health Journal, 2015. **19**(9): p. 2029-37.

172. Doherty, T., Chopra, M., Nsibande, D. and Mngoma, D., *Improving the coverage of the PMTCT programme through a participatory quality improvement intervention in South Africa.* BMC Public Health, 2009. **9**: p. 406.

173. Tomlinson, M., Doherty, T., Ijumba, P., Jackson, D., Lawn, J., Persson, L.A., et al., *Goodstart: a cluster randomised effectiveness trial of an integrated, community-based package for maternal and newborn care, with prevention of mother-to-child transmission of HIV in a South African township.* Tropical Medicine & International Health, 2014. **19**(3): p. 256-266.

174. Nsibande, D., Doherty, T., Ijumba, P., Tomlinson, M., Jackson, D., Sanders, D., et al., *Assessment of the uptake of neonatal and young infant referrals by community health workers to public health facilities in an urban informal settlement, KwaZulu-Natal, South Africa.* BMC Health Services Research, 2013. **13**: p. 47.

175. Smith, S.J., Nimmo, C., Fredlund, V. and Moodley, P., *Early infant diagnosis of HIV and fast initiation of anti-retroviral therapy in a rural African setting: how well are we doing?* Paediatrics and International Child Health, 2014. **34**(3): p. 203-7.

176. Fatti, G., Shaikh, N., Eley, B., Jackson, D. and Grimwood, A., *Adolescent and young pregnant women at increased risk of mother-to-child transmission of HIV and poorer maternal and infant health outcomes: A cohort study at public facilities in the Nelson Mandela Bay Metropolitan district, Eastern Cape, South Africa.* South African Medical Journal, 2014. **104**(12): p. 874-80.

177. Lilian, R.R., Kalk, E., Technau, K.-G. and Sherman, G.G., *Birth Diagnosis of HIV Infection in Infants to Reduce Infant Mortality and Monitor for Elimination of Mother-to-child Transmission.* The Pediatric Infectious Disease Journal, 2013. **32**(10): p. 1080-1085.

178. Chetty, T., Knight, S., Giddy, J., Crankshaw, T.L., Butler, L.M. and Newell, M.L., *A retrospective study of Human Immunodeficiency Virus transmission, mortality and loss to follow-up among infants in the first 18 months of life in a prevention of mother-to-child*

transmission programme in an urban hospital in KwaZulu-Natal, South Africa. BMC Pediatrics, 2012. **12**: p. 146.

179. Geddes, R., Giddy, J., Butler, L.M., Van Wyk, E., Crankshaw, T., Esterhuizen, T.M., et al., *Dual and triple therapy to prevent mother-to-child transmission of HIV in a resource-limited setting - lessons from a South African programme.* South African Medical Journal, 2011. **101**(9): p. 651-4.

180. Mazanderani, A., MacLeod, W., Bor, J. and Sherman, G. *Age at HIV diagnosis within South Africa's early infant diagnosis program, 2010-2015*. in *CROI 2018*. 2018. Boston, USA.

181. Mazanderani, A.H., Moyo, F. and Sherman, G.G., *Missed diagnostic opportunities within South Africa's early infant diagnosis program, 2010–2015.* PLOS One, 2017. **12**(5): p. e0177173.

182. Hsiao, N.-Y., Stinson, K. and Myer, L., *Linkage of HIV-Infected Infants from Diagnosis to Antiretroviral Therapy Services across the Western Cape, South Africa.* PLOS One, 2013. **8**(2): p. e55308.

183. Teasdale, C.A., Yuengling, K.A., Mutiti, A., Arpadi, S., Nxele, M., Pepeta, L., et al., *Delays in fast track antiretroviral therapy initiation and reasons for not starting treatment among eligible children in Eastern Cape, South Africa.* AIDS, 2019. **33**(13): p. 2099-2101.

184. Technau, K.-G., Strehlau, R., Patel, F., Shiau, S., Burke, M., Conradie, M., et al., *12-month outcomes of HIV-infected infants identified at birth at one maternity site in Johannesburg, South Africa: an observational cohort study.* The Lancet HIV, 2018. **5**(12): p. e706-e714.

185. Abrams, E.J., Woldesenbet, S., Soares Silva, J., Coovadia, A., Black, V., Technau, K.G., et al., *Despite Access to Antiretrovirals for Prevention and Treatment, High Rates of Mortality Persist Among HIV-infected Infants and Young Children.* The Pediatric Infectious Disease Journal, 2017. **36**(6): p. 595-601.

186. Mathivha, E., Olorunju, S., Jackson, D., Dinh, T.H., du Plessis, N. and Goga, A., *Uptake of care and treatment amongst a national cohort of HIV positive infants diagnosed at primary care level, South Africa.* BMC Infectious Diseases, 2019. **19**(Suppl 1): p. 790.

187. Teasdale, C.A., Sogaula, N., Yuengling, K.A., Wang, C., Mutiti, A., Arpadi, S., et al., *HIV viral suppression and longevity among a cohort of children initiating antiretroviral therapy in Eastern Cape, South Africa.* Journal of Acquired Immune Deficiency Syndromes, 2018. **21**(8): p. e25168.

188. Shiau, S., Strehlau, R., Technau, K.G., Patel, F., Arpadi, S.M., Coovadia, A., et al., *Early age at start of antiretroviral therapy associated with better virologic control after initial suppression in HIV-infected infants.* AIDS, 2017. **31**(3): p. 355-364.

189. Moyo, F., Haeri Mazanderani, A., Bhardwaj, S., Mhlongo, O., Smith, B., Ngoma, K., et al., *Constructing a treatment cascade from routine laboratory data for HIV-PCR positive children in two districts in South Africa*, in *10th Workshop on HIV Pediatrics*. 2018: Amsterdam, The Netherlands.

190. Osler, M., Hilderbrand, K., Hennessey, C., Arendse, J., Goemaere, E., Ford, N., et al., *A three-tier framework for monitoring antiretroviral therapy in high HIV burden settings.* Journal of the International AIDS Society, 2014. **17**: p. 18908.

191. Takuva, S., Brown, A.E., Pillay, Y., Delpech, V. and Puren, A.J., *The continuum of HIV care in South Africa: implications for achieving the second and third UNAIDS 90-90-90 targets.* AIDS, 2017. **31**(4): p. 545-552.

192. Mehta, U., Heekes, A., Kalk, E. and Boulle, A., *Assessing the value of Western Cape Provincial Government health administrative data and electronic pharmacy records in ascertaining medicine use during pregnancy.* South African Medical Journal, 2018. **108**(5): p. 439-443.

193. Statistics South Africa, *Recorded Live Births*. 2016, Statistics South Africa: Pretoria.

194. Dusetzina, S.B., Tyree, S., Meyer, A.M., Meyer, A., Green, L. and Carpenter, W.R., in *Linking Data for Health Services Research: A Framework and Instructional Guide*. 2014: Rockville (MD).

195. Zhu, Y., Matsuyama, Y., Ohashi, Y. and Setoguchi, S., *When to conduct probabilistic linkage vs. deterministic linkage? A simulation study.* Journal of Biomedical Informatics, 2015. **56**: p. 80-86.

196. Jaro, M.A., *Probabilistic linkage of large public health data files.* Statistics in Medicine, 1995. **14**(5-7): p. 491-498.

197. DuVall, S.L., Kerber, R.A. and Thomas, A., *Extending the Fellegi-Sunter probabilistic record linkage method for approximate field comparators.* Journal of Biomedical Informatics, 2010. **43**(1): p. 24-30.

198. Newcombe, H.B. and Kennedy, J.M., *Record linkage: making maximum use of the discriminating power of identifying information.* Communications of the ACM, 1962. **5**(11): p. 563-566.

199. Fellegi, I.P. and Sunter, A.B., *A theory for record linkage.* Journal of the American Statistical Association, 1969. **64**(328): p. 1183-1210.

200. Hill, S., Atkinson, J. and Blakely, T., *Anonymous record linkage of census and mortality records: 1981, 1986, 1991, 1996 census cohorts*. 2002, Department of Public Health, Wellington School of Medicine and Health Sciences.

201. Dellicour, S., Brasseur, P., Thorn, P., Gaye, O., Olliaro, P., Badiane, M., et al., *Probabilistic record linkage for monitoring the safety of artemisinin-based combination therapy in the first trimester of pregnancy in Senegal.* Drug Safety, 2013. **36**(7): p. 505-513.

202. Tromp, M., Reitsma, J.B., Ravelli, A.C., Méray, N. and Bonsel, G.J. *Record linkage: making the most out of errors in linking variables*. in *AMIA Annual Symposium Proceedings*. 2006. American Medical Informatics Association.

203. Winkler, W.E., *String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage.* 1990.

204. Navarro, G., *A guided tour to approximate string matching.* ACM Computing Surveys (CSUR), 2001. **33**(1): p. 31-88.

205. Kariminia, A., Butler, T.G., Corben, S.P., Levy, M.H., Grant, L., Kaldor, J.M., et al., *Extreme cause-specific mortality in a cohort of adult prisoners—1988 to 2002: a data-linkage study.* International Journal of Epidemiology, 2006. **36**(2): p. 310-316.

206. Pacheco, A.G., Saraceni, V., Tuboi, S.H., Moulton, L.H., Chaisson, R.E., Cavalcante, S.C., et al., *Validation of a hierarchical deterministic record-linkage algorithm using data from 2 different cohorts of human immunodeficiency virus-infected persons and mortality databases in Brazil.* American Journal of Epidemiology, 2008. **168**(11): p. 1326-1332.

207. Russell, R.C., *US Patent 1261167 A*. 1918.

208. Grannis, S.J., Overhage, J.M., Hui, S. and McDonald, C.J., *Analysis of a probabilistic record linkage technique without human review.* AMIA Annual Symposium Proceedings, 2003: p. 259-63.

209. Statistics South Africa, *Community survey 2016 - Provinces at a glance*. 2016, Statistics South Africa: Pretoria.

210. Statistics South Africa, *Recorded Live Births*. 2010, Statistics South Africa: Pretoria.

211. Statistics South Africa, *Recorded Live Births*. 2011, Statistics South Africa: Pretoria.

212. Statistics South Africa, *Recorded Live Births*. 2012, Statistics South Africa: Pretoria.

213. Statistics South Africa, *Recorded Live Births*. 2013, Statistics South Africa: Pretoria.

214. Statistics South Africa, *Recorded Live Births*. 2014, Statistics South Africa: Pretoria.

215. Statistics South Africa, *Recorded Live Births*. 2015, Statistics South Africa: Pretoria.

216. National Department of Health of South Africa, *The 2015 National Antenatal Sentinel HIV & Syphilis Survey Report*. October 2017: Pretoria.

217. Ferguson Jr, T., Peterson, E.D., Coombs, L.P., Eiken, M.C., Carey, M.L., Grover, F.L., et al., *Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: A randomized controlled trial.* JAMA, 2003. **290**(1): p. 49-56.

218. Chetty, T., Yapa, H.M.N., Herbst, C., Geldsetzer, P., Naidu, K.K., De Neve, J.W., et al., *The MONARCH intervention to enhance the quality of antenatal and postnatal primary health*

services in rural South Africa: protocol for a stepped-wedge cluster-randomised controlled trial. BMC Health Services Research, 2018. **18**(1): p. 625.

219. South African Government. *Minister Botha launches Road to Health Booklet*. 2011; Available from: www.gov.za/minister-botha-launches-road-health-booklet. Accessed on 10th September 2018.

220. Woldesenbet, S.A., Jackson, D., Goga, A.E., Crowley, S., Doherty, T., Mogashoa, M.M., et al., *Missed Opportunities for Early Infant HIV Diagnosis: Results of A National Study in South Africa.* Journal of Acquired Immune Deficiency Syndromes, 2015. **68**(3): p. e26-e32.

221. Office of the Rights of the Child (The Presidency Republic of South Africa), *Situational Analysis of Children in South Africa*. 2009, The Presidency: Pretoria.

222. Garenne, M., Collinson, M.A., Kabudula, C.W., Gomez-Olive, F.X., Kahn, K. and Tollman, S., *Completeness of birth and death registration in a rural area of South Africa: the Agincourt health and demographic surveillance, 1992-2014.* Global Health Action, 2016. **9**: p. 32795.

223. Massyn, N., Day, C., Peer, N., Padarath, A., Barron, P. and English, R., *District Health Barometer 2013/14*. 2014, Health Systems Trust: Durban.

224. Naidoo, H., Avenant, T. and Goga, A., *Completeness of the Road-to-Health Booklet and Road-to-Health Card: Results of cross-sectional surveillance at a provincial tertiary hospital.* Southern African Journal of HIV Medicine, 2018. **19**(1): p. 1-10.

225. UNAIDS. *Country Factsheets - South Africa*. 2017; Available from: http://www.unaids.org/en/regionscountries/countries/southafrica (Accessed on: 14th November 2018).

226. Bourne, D.E., Thompson, M., Brody, L.L., Cotton, M., Draper, B., Laubscher, R., et al., *Emergence of a peak in early infant mortality due to HIV/AIDS in South Africa.* AIDS, 2009. **23**(1): p. 101-6.

227. Newell, M.L., Coovadia, H., Cortina-Borja, M., Rollins, N., Gaillard, P. and Dabis, F., *Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis.* Lancet, 2004. **364**(9441): p. 1236-43.

228. Dobson, A.J. and Barnett, A.G., *An introduction to generalized linear models*. 2008: Chapman and Hall/CRC.

229. Cramer, H., *Mathematical Methods of Statistics*. Princeton U. Press, Princeton. 1946. 500.

230. Donders, A.R.T., van der Heijden, G.J.M.G., Stijnen, T. and Moons, K.G.M., *Review: A gentle introduction to imputation of missing values.* Journal of Clinical Epidemiology, 2006. **59**(10): p. 1087-1091.

231. Hughes, R.A., Heron, J., Sterne, J.A.C. and Tilling, K., *Accounting for missing data in statistical analyses: multiple imputation is not always the answer.* International Journal of Epidemiology, 2019. **48**(4): p. 1294-1304.

232. Bartlett, J.W., Harel, O. and Carpenter, J.R., *Asymptotically Unbiased Estimation of Exposure Odds Ratios in Complete Records Logistic Regression.* American Journal of Epidemiology, 2015. **182**(8): p. 730-6.

233. Menard, S.W., *Logistic regression: from introductory to advanced concepts and applications*. 2010, Thousand Oaks, Calif. London: SAGE.

234. Hedeker, D., du Toit, S.H.C., Demirtas, H. and Gibbons, R.D., *A note on marginalization of regression parameters from mixed models of binary outcomes.* Biometrics, 2018. **74**(1): p. 354-361.

235. Heinze, G., Wallisch, C. and Dunkler, D., *Variable selection - A review and recommendations for the practicing statistician.* Biometrical Journal. Biometrische Zeitschrift, 2018. **60**(3): p. 431-449.

236. Hernán, M.A., Hernández-Díaz, S., Werler, M.M. and Mitchell, A.A., *Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology.* American Journal of Epidemiology, 2002. **155**(2): p. 176-184.

237. Victora, C.G., Huttly, S.R., Fuchs, S.C. and Olinto, M.T., *The role of conceptual frameworks in epidemiological analysis: a hierarchical approach.* International Journal of Epidemiology, 1997. **26**(1): p. 224-7.

238. Hosmer Jr, D.W., Lemeshow, S. and Sturdivant, R.X., *Applied Logistic Regression*. Vol. 398. 2013: John Wiley & Sons.

239. Harrell Jr, F.E., *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2015: Springer.

240. Cui, J., *QIC program and model selection in GEE analyses.* Stata Journal, 2007. **7**(2): p. 209-220.

241. Pan, W., *Akaike's information criterion in generalized estimating equations.* Biometrics, 2001. **57**(1): p. 120-5.

242. Goggin, K., Wexler, C., Nazir, N., Staggs, V.S., Gautney, B., Okoth, V., et al., *Predictors of infant age at enrollment in early infant diagnosis services in Kenya.* AIDS and Behavior, 2016. **20**(9): p. 2141-2150.

243. Hampanda, K.M., Nimz, A.M. and Abuogi, L.L., *Barriers to uptake of early infant HIV testing in Zambia: the role of intimate partner violence and HIV status disclosure within couples.* AIDS Research and Therapy, 2017. **14**(1): p. 17.

244. Peltzer, K. and Mlambo, G., *Factors determining HIV viral testing of infants in the context of mother-to-child transmission.* Acta Paediatrica, 2010. **99**(4): p. 590-596.

245. Baker, L., *The South African expanded programme on immunisation schedule : vaccinology.* Professional Nursing Today, 2010. **14**(4): p. 19-22.

246. World Health Organization. *WHO vaccine-preventable diseases: monitoring system*. 2018; Available from: http://apps.who.int/immunization_monitoring/globalsummary/countries?countrycriteria%5Bcountry%5D%5B%5D=ZAF (Last accessed: 10th September 2018).

247. Massyn, N., Peer, N., English, R., Padarath, A., Barron, P. and Day, C., *District Health Barometer 2015/16*. 2016, Health Systems Trust: Durban.

248. Dunning, L., Francke, J.A., Mallampati, D., MacLean, R.L., Penazzato, M., Hou, T., et al., *The value of confirmatory testing in early infant HIV diagnosis programmes in South Africa: A cost-effectiveness analysis.* PLOS Medicine, 2017. **14**(11): p. e1002446.

249. Fidler, K.J., Foster, C., Lim, E.J., Patel, A., Welch, S., Menson, E., et al., *Reactivity of Routine HIV Antibody Tests in Children With Perinatally Acquired HIV-1 in England: Cross-sectional Analysis.* The Pediatric Infectious Disease Journal, 2019. **38**(2): p. 146-148.

250. Mutambo, C. and Hlongwana, K., *Healthcare Workers' Perspectives on the Barriers to Providing HIV Services to Children in Sub-Saharan Africa.* AIDS Research and Treatment, 2019. **2019**: p. 8056382-8056382.

251. Ndirangu, J., Bland, R., Barnighausen, T. and Newell, M.L., *Validating child vaccination status in a demographic surveillance system using data from a clinical cohort study: evidence from rural South Africa.* BMC Public Health, 2011. **11**: p. 372.

252. Ho, C.P., Yeh, J.I., Wen, S.H. and Lee, T.J., *Associations among medication regimen complexity, medical specialty, and medication possession ratio in newly diagnosed hypertensive patients: A population-based study.* Medicine (Baltimore), 2017. **96**(45): p. e8497.

253. Ford, N., Vitoria, M., Doherty, M. and Gray, A., *Candidates for inclusion in a universal antiretroviral regimen: are lamivudine and emtricitabine interchangeable?* Current Opinion in HIV and AIDS, 2017. **12**(4): p. 334-338.

254. Gesesew, H.A., Ward, P., Woldemichael, K. and Mwanri, L., *Late presentation for HIV care in Southwest Ethiopia in 2003-2015: prevalence, trend, outcomes and risk factors.* BMC Infectious Diseases, 2018. **18**(1): p. 59-59.

255. Gaede, B. and Versteeg, M., *The State of the Right to Health in Rural South Africa.* South African Health Review, 2011. **2011**(1): p. 99-106.

256. Tagarro, A., Chan, M., Zangari, P., Ferns, B., Foster, C., De Rossi, A., et al., *Early and Highly Suppressive Antiretroviral Therapy Are Main Factors Associated With Low Viral Reservoir in European Perinatally HIV-Infected Children.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2018. **79**(2): p. 269-276.

257. Médecins Sans Frontières, *Untangling the Web of Antiretroviral Price Reductions, 18th Edition*. 2016: Geneva, Switzerland.

258. Lilian, R.R., Mutasa, B., Railton, J., Mongwe, W., McIntyre, J.A., Struthers, H.E., et al., *A 10-year cohort analysis of routine paediatric ART data in a rural South African setting.* Epidemiology and Infection, 2017. **145**(1): p. 170-180.

259. Pascoe, S., Fox, M., Huber, A., Murphy, J., Phokojoe, M., Gorgens, M., et al., *Gaps in care or data issues? The challenges to reaching the 90-90-90 targets in South Africa.*, in *9th IAS Conference on HIV Science*. 2017: Paris, France.

260. Moorhouse, M., Conradie, F. and Venter, F., *What is the role of CD4 count in a large public health antiretroviral programme?* Southern African Journal of HIV Medicine, 2016. **17**(1): p. 1-3.

261. Porter, M., Davies, M.A., Mapani, M.K., Rabie, H., Phiri, S., Nuttall, J., et al., *Outcomes of Infants Starting Antiretroviral Therapy in Southern Africa, 2004-2012.* Journal of Acquired Immune Deficiency Syndromes, 2015. **69**(5): p. 593-601.

262. Jiamsakul, A., Kariminia, A., Althoff, K.N., Cesar, C., Cortes, C.P., Davies, M.A., et al., *HIV Viral Load Suppression in Adults and Children Receiving Antiretroviral Therapy-Results From the IeDEA Collaboration.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2017. **76**(3): p. 319-329.

263. Hwang, B., Shroufi, A., Gils, T., Steele, S.J., Grimsrud, A., Boulle, A., et al., *Stock-outs of antiretroviral and tuberculosis medicines in South Africa: A national cross-sectional survey.* PLoS One, 2019. **14**(3): p. e0212405.

264. Penazzato, M., Prendergast, A.J., Muhe, L.M., Tindyebwa, D. and Abrams, E., *Optimisation of antiretroviral therapy in HIV-infected children under 3 years of age.* Cochrane Database of Systematic Reviews, 2014(5).

265. Bonawitz, R., Brennan, A.T., Long, L., Heeren, T., Maskew, M., Sanne, I., et al., *Regimen durability in HIV-infected children and adolescents initiating first-line antiretroviral therapy in a large public sector HIV cohort in South Africa.* Tropical Medicine and International Health, 2018. **23**(6): p. 650-660.

266. Decloedt, E.H., McIlleron, H., Smith, P., Merry, C., Orrell, C. and Maartens, G., *Pharmacokinetics of lopinavir in HIV-infected adults receiving rifampin with adjusted doses of lopinavir-ritonavir tablets.* Antimicrobial Agents and Chemotherapy, 2011. **55**(7): p. 3195-200.

267. Collins, I.J., Wools-Kaloustian, K., Goodall, R., Smith, C., Abrams, E.J., Ben-Farhat, J., et al., *Incidence of switching to second-line antiretroviral therapy and associated factors in children with HIV: an international cohort collaboration.* The Lancet HIV, 2019. **6**(2): p. e105-e115.

268. Brinkhof, M.W., Pujades-Rodriguez, M. and Egger, M., *Mortality of patients lost to follow-up in antiretroviral treatment programmes in resource-limited settings: systematic review and meta-analysis.* PLoS One, 2009. **4**(6): p. e5790.

269. Johnson, L.F., Dorrington, R.E. and Moolla, H., *Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa.* South African Journal of HIV Medicine, 2017. **18**(1): p. 694.

270. Spectrum System of Policy Models, *Spectrum Manual*. 2019.

271. Goga, A.E., Jackson, D.J., Lombard, C., Ramokolo, V., Ngandu, N., Sherman, G., et al., *Highest risk of mother-to-child transmission of HIV or death in the first 6 months postpartum: Results from 18 month follow-up of an HIV-exposed national cohort, South Africa* in *AIDS 2016*. 2016: Durban, South Africa.

272. Goga, A.E., Dinh, T.-H., Jackson, D.J., Lombard, C.J., Puren, A., Sherman, G., et al., *Population-level effectiveness of PMTCT Option A on early mother-to-child (MTCT) transmission of HIV in South Africa: implications for eliminating MTCT.* Journal of Global Health, 2016. **6**(2): p. 020405-020405.

273. Iwuji, C.C., Orne-Gliemann, J., Larmarange, J., Balestre, E., Thiebaut, R., Tanser, F., et al., *Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial.* The Lancet HIV, 2017.

274. Ford, C., Chibwesha, C.J., Winston, J., Jacobs, C., Lubeya, M.K., Musonda, P., et al., *Women's decision-making and uptake of services to prevent mother-to-child HIV transmission in Zambia.* AIDS Care, 2018. **30**(4): p. 426-434.

275. Ibrahim, M., Maswabi, K., Ajibola, G., Moyo, S., Hughes, M.D., Batlang, O., et al., *Targeted HIV testing at birth supported by low and predictable mother-to-child transmission risk in Botswana.* Journal of the International AIDS Society, 2018. **21**(5): p. e25111.

276. Sandbulte, M.R., Gautney, B.J., Maloba, M., Wexler, C., Brown, M., Mabachi, N., et al., *Infant HIV testing at birth using point-of-care and conventional HIV DNA PCR: an implementation feasibility pilot study in Kenya.* Pilot and Feasibility Studies, 2019. **5**(1): p. 18.

277. Horwood, C., Voce, A., Vermaak, K., Rollins, N. and Qazi, S., *Routine checks for HIV in children attending primary health care facilities in South Africa: Attitudes of nurses and child caregivers.* Social Science & Medicine, 2010. **70**(2): p. 313-320.

278. Kim, M., Wanless, R.S., Ahmed, S., Mhango, J., Damba, D., Kayabu, A., et al. *Multi-month prescription of antiretroviral therapy and its feasibility: experiences from the Baylor International Pediatric AIDS initiative (BIPAI) in six southern African countries*. in *9th IAS Conference on HIV Science*. 2017. Paris, France.

279. Chetty, T., Newell, M.L., Thorne, C. and Coutsoudis, A., *Viraemia before, during and after pregnancy in HIV-infected women on antiretroviral therapy in rural KwaZulu-Natal, South Africa, 2010-2015.* Tropical Medicine & International Health, 2018. **23**(1): p. 79-91.

280. Larsen, A., Magasana, V., Dinh, T.H., Ngandu, N., Lombard, C., Cheyip, M., et al., *Longitudinal adherence to maternal antiretroviral therapy and infant Nevirapine prophylaxis from 6 weeks to 18 months postpartum amongst a cohort of mothers and infants in South Africa.* BMC Infectious Diseases, 2019. **19**(Suppl 1): p. 789.

281. Harklerode, R., Schwarcz, S., Hargreaves, J., Boulle, A., Todd, J., Xueref, S., et al., *Feasibility of Establishing HIV Case-Based Surveillance to Measure Progress Along the Health Sector Cascade: Situational Assessments in Tanzania, South Africa, and Kenya.* JMIR Public Health and Surveillance, 2017. **3**(3): p. e44-e44.

282. Organization, W.H., *Consolidated guidelines on person-centred HIV patient monitoring and case surveillance.* 2017.

283. Boulle, A., Heekes, A., Tiffin, N., Smith, M., Mutemaringa, T., Zinyakatira, N., et al., *Data Centre Profile: The Provincial Health Data Centre of the Western Cape Province, South Africa.* International Journal of Population Data Science, 2019. **4**(2).

284. Heekes, A., Tiffin, N., Dane, P., Mutemaringa, T., Smith, M., Zinyakatira, N., 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 Global Health, 2018. **3**(Suppl 2): p. e000565.

285. Baisley, K., Seeley, J., Siedner, M., Koole, K., Matthews, P., Tanser, F., et al., *Findings from home-based HIV testing and facilitated linkage after scale-up of test and treat in rural South Africa: young people still missing.* HIV Medicine, 2019. **20**(10): p. 704-708.

286. Nicol, E., Dudley, L. and Bradshaw, D., *Assessing the quality of routine data for the prevention of mother-to-child transmission of HIV: An analytical observational study in two health districts with high HIV prevalence in South Africa.* International Journal of Medical Informatics, 2016. **95**: p. 60-70.

287. Nicol, E., Bradshaw, D., Uwimana-Nicol, J. and Dudley, L., *Perceptions about data-informed decisions: an assessment of information-use in high HIV-prevalence settings in South Africa.* BMC Health Services Research, 2017. **17**(Suppl 2): p. 765.

288. Goga, A., Singh, Y., Jackson, D., Mukungunugwa, S., Wafula, R., Eliya, M., et al., *How are countries in sub-Saharan African monitoring the impact of programmes to prevent vertical transmission of HIV?* British Medical Journal, 2019. **364**: p. l660.

289. Katurura, M.C. and Cilliers, L., *Electronic health record system in the public health care sector of South Africa: A systematic literature review.* African Journal of Primary Health Care & Family Medicine, 2018. **10**(1): p. 1-8.

290. Beck, E.J., Shields, J.M., 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?* Global Health Action, 2018. **11**(1): p. 1440782-1440782.

291. Lopetegui, M., Yen, P.-Y., Lai, A., Jeffries, J., Embi, P. and Payne, P., *Time motion studies in healthcare: What are we talking about?* Journal of Biomedical Informatics, 2014. **49**: p. 292-299.

292. Spooner, E., Govender, K., Reddy, T., Ramjee, G., Mbadi, N., Singh, S., et al., *Point-of-care HIV testing best practice for early infant diagnosis: an implementation study.* BMC Public Health, 2019. **19**(1): p. 731.

293. Jacob, N., Rice, B., Kalk, E., Heekes, A., Morgan, J., Hargreaves, J., et al., *Digitising point of care HIV test results to accurately measure, and improve performance towards, the UNAIDS 90-90-90 targets.* medRxiv, 2019: p. 19012302.

294. Dorward, J., Drain, P.K. and Garrett, N., *Point-of-care viral load testing and differentiated HIV care.* Lancet HIV, 2018. **5**(1): p. e8-e9.

295. Dunning, L., Penazzato, M., Soeteman, D., Gandhi, A., Phillips, A., Dugdale, C., et al., *The cost-effectiveness of routine HIV screening and testing at infant immunization visits in Côte d'Ivoire, South Africa, and Zimbabwe*, in *11th International Workshop on HIV Pediatrics*. 2019: Mexico City, Mexico.