Who died, where, when and why?

An investigation of HIV-related mortality in rural South Africa

Paul Mee
2015
DEDICATION

To Kathleen, Daniel and Vivienne
for giving so much more than you know.

To all the courageous people of South Africa
who refuse to let the spirit of Ubuntu be extinguished.
You are my inspiration.
CONTENTS

ABSTRACT ........................................................................................................................................i

ORIGINAL PAPERS ..........................................................................................................................ii

ABBREVIATIONS .............................................................................................................................iii

FOREWORD ........................................................................................................................................1

INTRODUCTION ...............................................................................................................................2
  Conceptual Framework and summary of papers comprising this work ............................................3

BACKGROUND ....................................................................................................................................5
  South African history ...........................................................................................................................5
  The development of the HIV epidemic ...............................................................................................6
  Healthcare provision in South Africa .................................................................................................9
  Political response to the AIDS epidemic in South Africa .................................................................10
  Structural and social determinants of HIV ......................................................................................13
  Migration ............................................................................................................................................17
  Poverty and education .........................................................................................................................20
  Gender and age ..................................................................................................................................25
  Physical barriers to accessing HIV services .....................................................................................26
  The impact of health related belief systems on the HIV/AIDS epidemic in Sub-Saharan Africa ....29
  The history of demographic surveillance .........................................................................................31
  Verbal autopsy ...................................................................................................................................35

METHODS ..........................................................................................................................................39
  Study setting and the establishment of the Agincourt HDSS study site ........................................39
  In-migration to the Agincourt region from Mozambique .................................................................40
  Characteristics of the Agincourt HDSS study site ..........................................................................44
  Spatial Analysis and its relevance to this study ................................................................................47
  Micro-epidemic analysis ...................................................................................................................48

RESULTS ............................................................................................................................................50
ABSTRACT

Background

South Africa has experienced the most severe consequences of the HIV/AIDS pandemic. Every community has been affected in some way, many experiencing huge increases in mortality, particularly before antiretroviral therapies (ART) were readily available. However, the micro-level understanding of the HIV epidemic in South Africa is weak, because of a lack of detailed data for most of the population. This thesis is based on detailed individual follow-up in the Agincourt Health and Demographic Surveillance Site (HDSS) located in the Agincourt sub-district of Mpumalanga Province and investigates micro-level determinants of HIV epidemiology and the impact of treatment provided.

Methods

The Agincourt HDSS has followed a geographically defined population since 1992, approximately the time when the HIV/AIDS epidemic first became apparent. This population-based surveillance has included capturing details of all deaths, with cause of death determined by verbal autopsy, as well as the geographical location of individual households within the overall Agincourt area. Background information on the roll-out of ART over time was also recorded.

Results

A comparison immediately before and after the major roll-out of ART showed a substantial decrease in HIV-related mortality, greater in some local communities within the area than others. Individual determinants associated with a decreased risk of HIV/AIDS mortality included proximity to ART services, as well as being female, younger, and in higher socioeconomic and educational strata. There was a decrease in the use of traditional healthcare sources and an increase in the use of biomedical healthcare amongst those dying of HIV/AIDS between periods before and after the roll-out of ART.

Conclusions

Understanding micro-level determinants of HIV/AIDS infection and mortality was very important in terms of characterising the overall epidemic in this community. This approach will enable public health interventions to be more effectively targeted towards those who need them most in the continuing evolution of the HIV/AIDS epidemic.
ORIGINAL PAPERS

This thesis is based on the following papers:

I
"The development of a localised HIV epidemic and the associated excess mortality burden in a rural area of South Africa"
(Submitted).

II

III
"Determinants of the risk of dying of HIV/AIDS in a rural South African community over the period of the decentralised roll-out of antiretroviral therapy: a longitudinal study."

IV

All published works are in open access journals allowing their reproduction in this thesis.
# ABBREVIATIONS

**AIDS** - Acquired Immunodeficiency Syndrome

**ANC** - African National Congress

**ART** - Antiretroviral Therapy

**AZT** - Azidothymidine (also known as Zidovudine)

**CCC** - Concordance Correlation Coefficient

**CCVA** - Computer Coded Verbal Autopsy

**CSMF** - Cause Specific Mortality Fraction

**CSW** - Commercial Sex Workers

**DHS** - Demographic and Health Surveys

**DRC** - Democratic Republic of Congo

**HDSS** - Health and Demographic Surveillance System

**HIV** - Human Immunodeficiency Virus

**HRS** - Household Registration System

**HSDU** - Health Systems Development Unit

**INDEPTH** - International Network of field sites with continuous Demographic Evaluation of Populations and Their Health

**IPV** - Intimate Partner Violence

**MDG** - Millennium Development Goals

**MSM** - Men who have Sex with Men

**NGO** - Non-Governmental Organisation

**PCVA** - Physician Coded Verbal Autopsy

**PHMRC** - Population Health Metrics Research Consortium

**PMTCT** - Prevention of Mother to Child Transmission

**PWID** - Person Who Injects Drugs

**RDBMS** - Relational Database Management System

**SIV** - Simian Immunodeficiency Virus

**TAC** - Treatment Action Campaign

**TB** - Tuberculosis

**UNAIDS** - United Nations Agency on HIV/AIDS

**WHO** - World Health Organization
FOREWORD

At one level, this dissertation can be read as an analysis of the changes in patterns of mortality and health seeking behaviour associated with the HIV epidemic in a rural South African community. An alternative reading could however be that it is the story of how this community and their health providers responded to the unprecedented threat of the HIV/AIDS epidemic, which in the space of a few years developed from a few sporadic cases to become by far the leading cause of death. The nature of HIV/AIDS, preferentially targeting the youngest and most productive members of the population, brought huge social and economic burdens. The dramatic later reversal of the trajectory of the epidemic is a testimony to the resilience of this community in the face of adversity, to care for those who were sick and to take on board new understandings about how this disease could effectively be managed and treated.

We are, in 2015, at a pivotal moment in the global battle to respond to HIV/AIDS. Despite remarkable scientific breakthroughs, effective vaccines and cures are still some way off. Despite this, using our current proven approaches to prevention and treatment we have the potential to reduce the current global epidemic to endemic levels over the next fifteen years. This will require increases in, and more effective targeting of, the resources committed to tackling the disease. This ultimately depends on the availability of data characterising the HIV epidemic at both macro and micro levels and the development of analytical methods to enable health planners to make more effective use of those data. The challenges are great, especially given the current efforts required to address global issues such as of population growth, climate change and societal upheaval, all of which make increasing demands on the resources of governments and donors of aid. The economic and human costs of failure however will be far greater than the investment needed now to respond.

Scientific progress is by its nature a collaborative effort. I hope that this work will make a contribution to this effort and provide ideas and inspiration for others.
INTRODUCTION

The world is now approximately thirty years into the HIV epidemic. When considering the burden of HIV/AIDS and the effect of the response to the epidemic it is useful and sobering to review the numbers, even with the inevitable risk of depersonalisation that comes with viewing the epidemic through a purely statistical lens.

In 2013 worldwide, there were 2.1 million adults and children newly infected with HIV, 1.5 million deaths were attributable to HIV/AIDS and 35 million people were infected with HIV. The vast majority of those affected lived in the countries of Sub-Saharan Africa where in 2013 25 million people were estimated to be living with HIV and 1.5 million were newly infected [1]. Looking at the relative burden in this region compared with the rest of the world the differences become even more dramatic. In Sub-Saharan African countries the prevalence of HIV in 2013 was 2.7 %, if we look at the worldwide figures excluding this region the prevalence was 0.2 % [1, 2]. Thus an individual living in Sub-Saharan Africa was roughly 17 times more likely to be infected with HIV than someone living in the rest of the world.

Focusing on South Africa the situation is even more dramatic; in 2013 6.3 million people out of a population of 54 million were living with HIV giving an overall HIV prevalence of 11.9 %. For adults aged 15 to 49 the prevalence was 19.1 % [1]. It is estimated that in 2013 there were 195,000 HIV related deaths and 341,000 new infections, thus in that year an additional 146,000 people were added to the total number of South Africans living with HIV [1]. In the short term at least this trend is likely to continue and the numbers living with HIV and ultimately requiring treatment, the vast majority within the South African national healthcare system, will continue to increase.

Recent studies indicate that with appropriate levels of care the life expectancy of those living with HIV is comparable to those uninfected. An HIV-infected individual in the USA and Canada receiving adequate and timely treatment can expect to live until their early 70s, a figure approaching that of the general population [3].

This work focuses on the changing patterns of the epidemic, risk factors associated with HIV/AIDS mortality and changes in health seeking behaviour among a rural population in the north east of South Africa. The ultimate goal is to add to the evidence base available to healthcare policy makers as they seek to improve the coverage and effectiveness of the HIV testing, treatment and prevention services in the country.
Conceptual Framework and summary of papers comprising this work

This dissertation comprises four individual manuscripts which are briefly described below.

Paper I is an analysis of the overall patterns of development of the epidemic from 1992, when the first cases were identified in the study area, to 2013. The paper aims to identify whether the epidemic was originally localised either in geographic space or in particular population sub-groups and then how it developed. Comparisons are made between the local epidemic profile and that of the national epidemic in South Africa in order to understand the generalisability of the patterns seen.

Paper II is a spatial analysis of the changes in rates of HIV/AIDS mortality over the period of the ART roll-out (2007 – 2010). Spatial data are aggregated by village and also by grid square. The latter gives a sub-village breakdown of the patterns of mortality change enabling a fine grained understanding of local patterns of change within the epidemic. Cluster analysis is carried out to identify local epidemic hotspots.

Paper III complements paper II and provides an analysis of the socio-demographic and geographic determinants which were associated with the risk of dying of HIV/AIDS over the same period. The goal of this is to understand which sub-groups of the population had an increased risk of dying of HIV/AIDS over the period of the ART roll-out.

Paper IV assesses the evidence for changes in health seeking behaviour amongst those dying of HIV/AIDS over a longer period (2003 – 2011). The emphasis is to assess evidence for the adoption of biomedical treatment as an alternative to traditional remedies once it became available.

The common themes running through these papers as they relate to the local and national HIV/AIDS epidemics are shown in the conceptual framework diagram in Figure 1.
Figure 1 Dissertation schematic showing the main themes running through the research and the constituent papers in which they are covered.
BACKGROUND

In order to gain an understanding of development of the HIV/AIDS epidemic in South Africa, the response to it from the national government and the subsequent effects in terms of patterns of mortality and morbidity, it is essential to understand the unique social and political history of the country. The next section briefly summarises this history describing the linked social and political development of the country and the provision of healthcare to its population.

South African history

The first European settlers in South Africa, predominantly of Dutch origin arrived in 1652 and established what became known as the Cape Colony. One consequence of this was a developing destruction, or forcing to the margins, of the indigenous peoples, the Khoi Khoi and San tribes, who had lived in the region for between 10,000 and 20,000 years. This occurred through territorial conflict and the introduction of previously unknown infectious diseases such as sexually transmitted infections. The Dutch East India Company controlled the area until the British occupation in 1806 which preceded a century of British rule. During this period groups of armed Afrikaner farmers (the Voortrekkers) began to migrate inland and occupy much of what constitutes modern South Africa. The indigenous population of the occupied lands were forced to either become tenants or were relocated to relatively impoverished rural areas. In 1910, following the Anglo-Boer war, the Union of South Africa was established combining the existing Afrikaner republics (Transvaal and Free State) and the British colonies (The Cape and Natal) [4].

A dramatic transformation of the social and economic structure of the country occurred with the discovery of diamonds in Kimberley in 1867 and gold in the Witwatersrand (the region around Johannesburg) in 1886. The subsequent development of the mining industry led to a huge demand for cheap male labour. To meet this need a variety of measures were introduced to force rural males to migrate to urban areas to work in the mines, with a consequent impact on the economies of the rural areas. The majority of the migrant workers lived in overcrowded insanitary slums and hostels in the urban areas, returning regularly to their families in the rural areas of the country. [4, 5]

In 1948 the National Party came to power in South Africa and over the subsequent years introduced the state policies of ‘Apartheid’, an Afrikaans word meaning literally ‘the state of being apart’. These policies sought to both solidify and expand the existing inequalities in order to benefit the white minority. Under the National Party government several pieces of legislation were enacted to give a legal framework to the controls on residence that had developed under the colonial era. Amongst the first of these was the Group Areas Act (1950), a series of pieces of legislation which designated which urban areas were allocated to the different ethnic groups and led to many of the black and coloured community being relocated to newly established townships, such as Soweto (south-west of Johannesburg) and the Cape Flats (close to Cape Town). Effectively these acts secured the prime areas close to the central
business districts for the white population of the country. The Bantu Homelands citizenship Act (1970) created a series of homelands or Bantustans for each of the ethnic groups within the country. The allocation of individuals to one of these groups was often arbitrary, failing to reflect the historical ethnic diversity of the peoples. The effect of this was that members of the black population only had a right of residence in their assigned homeland. These policies led in turn to a programme of forced removals. In the urban areas the residents of previously multiracial communities were relocated to newly established townships outside of the city. In many rural areas such forced removals served to consolidate the white ownership of the prime farmland. The black population were typically removed to newly established communities in the nearby area from where they could provide a source of labour [4, 6].

The continued and increasing need for a large male manual labour force coupled with the restrictions on where the black population were allowed to live led to a common pattern in which the families of the working men remained behind in their rural home whilst the working males spent the majority of their time living in the urban areas. The males would return to the rural community although these visits may have occurred as infrequently as once a year. Inevitably the men were forming sexual relationships, either transitory or more permanent in nature, in the cities [4, 7].

The ending of Apartheid came after the first democratic elections in South Africa in 1994, in which the African National Congress (ANC) came to power. Subsequently, although the legislative framework which enforced the societal fragmentation came to an end, the underlying structures were much slower to change. The patterns of high levels of labour migration and the associated disruption of family structures, unemployment and poverty, continue to the present day [8].

The development of the HIV epidemic

The syndrome which later came to be described as AIDS was first described in case studies of homosexual male patients in Los Angeles in the USA in June 1981, and further cases were subsequently reported in other US cities [9]. By 1983 other routes of infection were identified; transmission via heterosexual sex, infection from an infected mother to her child during childbirth and infection due to contaminated blood products [10, 11]. By 1984 the first cases of AIDS in Africa were seen. Piot et al. reported a case series of 38 HIV positive patients in Kinshasa, Zaire (now the Democratic Republic of Congo (DRC)). All of these were heterosexual and the females identified were younger than the males. There was a high prevalence of opportunistic infections reported amongst this group. This was described by the authors as a ‘new epidemiological setting for AIDS’ [12]. In the same year Van de Perre et al. reported on a 26 patient case series in neighbouring Rwanda; the males were all heterosexuals of relatively high socioeconomic position typically with multiple partners, and 4 of the 9 females identified were commercial sex workers [13]. Other reports subsequently began to emerge from this region and West Africa. The characteristics of the epidemic seen in these two early studies; the gender-age differentials, the heterosexual route of transmission, the high
prevalence in sex workers and the disproportionate impact on the more wealthy in the early stages of the epidemic all typify the sub-Saharan epidemic that was to subsequently develop. DeCock [9] writing in 2012 describes the emergence of HIV/AIDS in Africa as a “patchwork of epidemics moving through different groups and communities at different times”.

The rapid development of the HIV/AIDS epidemic in east Africa is shown in a series of cross-sectional surveys of HIV prevalence amongst female sex workers in Nairobi, Kenya. HIV prevalence in this group increased from 4 % in 1981 to 61 % in 1985 [14]. The authors of this study hypothesised three possible routes by which the virus could have been spread to this region; firstly via HIV infected commercial sex workers in Nairobi the majority of whom were from Tanzania, a neighbouring country already known to have high HIV prevalence, secondly via HIV infected male travellers from central Africa arriving in Nairobi and passing on the infection to sex workers there and finally by continuous transmission from those countries bordering the west of Kenya, with overland transport routes into the country e.g. Uganda, Rwanda and Burundi.

Molecular phylogenetic studies of wild chimpanzee and ape populations and analysis of human blood samples collected in the 1960s have helped us gain an understanding of both the origins of HIV and the routes of spread of the virus across Africa and throughout the world [15]. Apes and chimpanzees are hosts of the closely related Simian Immunodeficiency Virus (SIV). This is thought to have been transmitted to the human population in South East Cameroon in the early part of the 20th Century. Comparing the species specificity of the SIV and HIV subtypes showed that HIV-1, which predominates in eastern and Southern Africa, and HIV-2, a common subtype in Western African countries, are hosted by different species of primates. This indicates that at least two separate cross-species infectious events acted as precursors of the epidemic across Africa.

River based trade routes linked this region to the rapidly growing Zairean city of Leopoldville, now known as Kinshasa in the DRC. It was here that the virus first became established in the population and can thus be thought of as the cradle for the disease [15]. Further insights into the subsequent spread of the virus have come from analysing the different subtypes of HIV within the DRC and across Africa [16]. The internal land transport links within the DRC are poor, with a combination of river transport, paved, gravel and dirt roads making travel between the major population centres often very difficult. As result the populations of the peripheral cities are more likely to travel across the national borders into neighbouring countries than to travel internally within the country. Amongst HIV positive individuals in Kinshasa, which is in the west of the country, a mix of HIV subtypes similar to that in West African countries was seen. Lubumbashi in the south is close to the border with Zambia and hence linked to the paved road network in Zambia and Southern Africa; here HIV subtype B predominates which is the common subtype in Southern African countries. Kisangani in the east is connected by road and water (lake transport) to East African countries. In this region subtype A predominates as is the case in East African countries. As hypothesised by Piot et al. more than a decade earlier [14], Gray et al. [16] suggest that individuals migrating across the land borders such as commercial
sex workers, truck drivers and soldiers may have been mainly responsible for the diseases spread from its central African origins to the rest of the continent.

The first cases of HIV in South Africa were reported amongst a small number men who have sex with men (MSM) in 1982 [17]. It is thought that the introduction within this community was a result of individuals becoming infected when travelling overseas. Subsequently between 1982 and 1984 around 100 haemophiliacs contracted the virus through infected blood products [17]. In each of these groups the virus was subtype B. At this stage cases in the general population were rare. By 1988 there had been 166 patients with AIDS identified in South Africa of these 125 (75%) were among MSM, 24 (15%) among heterosexuals, 8 (5%) among haemophiliacs, 6 (4%) through blood transfusion and 3 (2%) paediatric cases. Of these 22 cases were in black South Africans amongst which14 (63.6 %) were from the Durban/Pietermaritzburg area of KwaZulu-Natal. Amongst the HIV infected black South Africans, the male to female ratio was 1:1 [18]. A survey on sera collected in 1985 in a rural community in KwaZulu-Natal identified no cases of HIV infection. By 1991 it was found that the prevalence of HIV in this same community was 2.2 % [19].

A series of repeated surveys of HIV prevalence amongst women attending antenatal clinics in South Africa demonstrates the subsequent rapid development of the epidemic in the country, the results are shown in Figure 2. For example in KwaZulu-Natal in the east the prevalence in this group was around 2 % in 1990 and had risen to around 35 % 10 years later and around 40 % by 2004, a level it stayed close to until around 2010. Other provinces show a similar pattern of a rapid increase and later plateau. However the peak prevalence showed a great deal of variation between the different provinces; the lowest level was in the Western Cape and Northern Cape at less than 20 %. KwaZulu-Natal on the eastern seaboard contains major ports such as Richards Bay and Durban, used for the transport of goods from other African countries. Collinson et al. [20] suggested that the sex trade associated with these sea ports provided the point of entry for the virus into the heterosexual population. Male labour migrants from rural KwaZulu-Natal mainly travelled to Gauteng, Durban, Empangeni and the ports on the eastern seaboard [21]. This may explain why the epidemic in KwaZulu-Natal preceded that in Mpumalanga and Gauteng. The lower levels of prevalence in the more isolated and relatively impoverished provinces such as Limpopo, the Northern Cape and Eastern Cape may be explained by their relative lack of connectedness through migration to the worst affected provinces. As will be discussed later the social conditions in the major urban areas also contributed to the development of the epidemic. The earlier more effective response of the healthcare system to the epidemic in the Western Cape may explain why the prevalence is much lower there than in other areas of the country.
Healthcare provision in South Africa

There is a history of fragmentation in health provision in South Africa. From as early as 1879 health facilities were racially segregated, and separate provision was established for curative and preventative services [4, 5].

In pioneering work to develop more effective models for the delivery of healthcare in the 1930s, Sidney and Emily Kark and colleagues established a network of community health centres in the Pholela region of Natal province [4, 23, 24]. Based on the evidence for the effectiveness of this model, the Gluckman commission proposed in 1944 that a chain of community health centres should be established across the country providing locally based primary care. However after the National Party came to power in 1948 these plans were not implemented. Under the Apartheid regime, provision was further fragmented, with each homeland having its own semi-autonomous health department. Spending per head on health services was significantly lower in the homeland areas than in the rest of the country. Prior to the 1960s mission hospitals provided the backbone of health provision in the homeland areas [4, 25].
Despite the inequities in health service provision under successive National Party governments, the post-Apartheid administration in South Africa inherited a health system that was relatively well funded by the standards of middle income countries, with 8.5% of the country’s GDP going into health. There was however, an emphasis on provision of care in the tertiary and academic hospitals with a low proportion of the total spend (11%) going into primary care. Their focus in developing the system was to follow the principles set out in the Alma Ata agreement [26] and the ideas set out by the Gluckman commission in the 1940s of shifting the focus to primary care. The aim of this was to redress historical inequities and ensure that the disadvantaged received essential healthcare [4]. The complex Apartheid era administrative structures were rationalised to create a system in which district health systems delivered primary care reporting to provincial health departments, with the national health department providing centralised coordination. There was a significant investment in the building and upgrading of clinics throughout the country and primary care was provided at no cost. Mass immunisation campaigns were launched and drug supply improved. The inequities in per-capita spending between provinces were also reduced [4].

However many challenges remained. Prior to 1994 very few black doctors had been trained. Also there was a lack of emphasis on primary healthcare provision in the training of doctors in the country. The new government brought in employment policies designed to provide job opportunities for those from previously disadvantaged racial groups. One consequence of this was that many experienced middle and senior managers left the public healthcare system. Due to this sudden rather than gradual transition there was a loss of management capacity in the healthcare sector. At the same time many skilled medical professionals chose to leave the public sector and move to the private health sector in South Africa or to positions overseas [27]. The consequence was that in the early stages of the development of the HIV epidemic in South Africa, the healthcare system suffered from structural challenges which limited its ability to respond as effectively to the epidemic as might otherwise have been the case.

**Political response to the AIDS epidemic in South Africa**

The Fourth International Conference on Health in Southern Africa was held in Maputo, Mozambique in 1990. The ANC together with non-governmental groups and health workers from South Africa and the surrounding areas participated in this meeting. The resulting ‘Maputo Convention’ acknowledged the urgent need to prioritise HIV prevention and to protect the human rights of those with the disease [28]. There was however a feeling that these positive initiatives were not matched by action and attitudinal change at the grass roots. There were limited efforts at education and prevention from the ruling National Party government, but these were undermined by the distrust felt by many towards this avowedly racist party [27]. During this period there was an increased climate of discrimination and stigma towards those with HIV.
In 1992 as the transition towards democracy gathered momentum, the South Africa National Department of Health working together with the ANC and others developed the National AIDS Coordinating Committee of South Africa national AIDS strategy. This proposed a multi-sectoral response to the developing epidemic. The document was adopted by the new ANC led government after the first democratic elections in 1994 and an immediate policy implementation was expected. However during the next few years, as the focus shifted towards enacting the democratic transition, implementing the response to HIV/AIDS became less of a priority alongside all of the competing challenges faced by the new government, and the response weakened [28].

Mary Crewe writing in 2000 reflects on the challenges over the subsequent period. Notable events included a significant rift between the NGO community and government over the AIDS education play Sarafina II commissioned by the Government, which became mired in accusations of financial mismanagement [29]. Subsequently the Government supported the testing of a supposed treatment for AIDS, Virodene, which proved to be an ineffective and toxic industrial solvent, rather than antiretroviral therapy (ART) drugs of proven efficacy [27, 29]. In 1998 the government refused to fund the provision of AZT for pregnant women to prevent mother-to-child transmission (PMTCT) of HIV [17]. The focus of the government continued to be on prevention rather than treatment. The Western Province, the only province in the country not run by the ANC, independently decided to introduce AZT for PMTCT into the public healthcare system in 1999 [30].

Following the second democratic elections in 1999 Thabo Mbeki and his health minister Manto Tshabala-Msimang began to promote unorthodox theories which questioned the link between HIV and AIDS and suggested that ART drugs were toxic. Mbeki proposed that poverty rather than HIV was the underlying cause of the epidemic and Tshabala-Msimang promoted nutritional interventions as opposed to ART for those with the disease [31]. Mbeki chose to convene an international advisory panel on HIV/AIDS comprising 16 so-called AIDS dissidents and 16 mainstream scientists rather than working with expert South African scientists and clinicians [32].

This period was also marked by the development of grass roots activist groups opposing Government policy on HIV/AIDS, notably the Treatment Action Campaign (TAC). TAC and others led a series of court cases against the government and in 2001 the South African Constitutional Court ruled that the government was obliged to provide nevirapine to all HIV positive pregnant women in South Africa to protect their babies from infection during childbirth [17]. In 2003 the Mbeki government agreed to provide ART for free in public health facilities. It has been estimated that if the decision had been taken 3 years earlier 330,000 lives would have been saved [33].

Mbeki’s second term as president was marked by his withdrawal from discussions on HIV whilst Tshabala-Msimang continued to undermine the HIV/AIDS treatment programme and
promote untested traditional remedies and vitamin supplements. She was eventually replaced as Minister of Health in 2006. In 2008 Thabo Mbeki was forced to resign and under the interim government of Kgalema Motlanthe and the subsequent election of Jacob Zuma, a much more positive approach emphasising the importance of prevention and treatment and accelerating ART roll-out across the country began. The National Strategic Plan of 2007-2011 led to a dramatic scaling up of ART provision across the country [34].

Subsequently the South African government faced a number of challenges in rolling out treatment throughout the country, these were particularly acute in rural areas. Supply side issues included a lack of political will to drive the process, the relatively low spending on public health relative to other sectors at the time and a lack of adequate infrastructure to process the patients needing to start treatment. Often the hospitals, where treatment was provided were operating at full capacity and the local clinics, to which in theory down-referral could occur, lacked the personnel and infrastructure to provide this service. There was also a shortage of doctors, who were required to initiate treatment, and again this was particularly acute in the rural areas of the country. From the demand side, the denialism of the link between HIV and AIDS coming from the government and the messages of mistrust around the efficacy and safety of antiretroviral treatments led many of those infected with HIV to seek alternative forms of care even when treatment was available [17, 25].

In recent years, significant progress has been made by the ANC government. By 2012 2.6 million people were on antiretroviral treatment in South Africa, 75 % of those living with HIV who were eligible [1]. However a number of significant challenges still remained. As a result of the success of the ART programme the numbers on treatment continued to increase with the resultant burden on the health services and the cost of providing medication. Also the HIV incidence rate remained essentially unchanged at around 1.7 % per annum from 2005 to 2012 indicating that there has been little change in sexual behaviour [22, 35].

**Structural and social determinants of HIV**

The following section gives some background and context to the risk factors for HIV mortality that have been investigated in this study.

It was recognised by Sidney Kark in 1949 that the social conditions in which individuals were living were largely responsible for the increase in prevalence of sexually transmitted infections [36]. This early study investigated syphilis infection, but many of the arguments made about the social causes apply equally to the HIV epidemic [23]. An important insight made by Kark was the importance of understanding the determinants of the development of an epidemic at both individual and societal levels, with an emphasis on the importance of the latter. In a commentary on Kark’s paper by Landon Myer in 2003 the author notes the parallels between the syphilis epidemic and the HIV epidemic [23]. In particular, the emphasis made on individual behaviour as the driver for the epidemic, rather than societal structure, comes partly from a belief that social and economic conditions are not amenable to intervention. The
parallels in terms of the underlying causes and the potential for a successful intervention between the syphilis epidemic of Kark’s time and the HIV epidemic are well illustrated by this quote from Kark’s paper:

“Without an understanding of the economic factors involved and the historical development of the vast social pathological changes brought about during the last 70 years, no treatment will save the spread of syphilis in South Africa. Treatment of individual personality disorder or attempts to inculcate a re-orientation towards a healthy sexual and family life cannot succeed in any but a few cases. The first line of treatment must be to remedy the unhealthy social relationships which have emerged as the inevitable result of masses of men leaving their homes every year.” [36]

These ideas were developed in a highly prescient paper published by Zwi and Cabral in 1991 before the South African HIV epidemic had become established [37]. In this they recognised that a combination of social, economic and political forces in the country at that time created what they described as a ‘high risk situation’. They describe the way in which factors such as unemployment, migration, rapid peri-urban settlement, poverty, landlessness and homelessness which were all highly prevalent in South Africa, particularly amongst the black population, could lead to societal disruption creating an environment in which there is a tendency towards higher risk forms of sexual behaviour. They argue for a focus on the underlying determinants of health as a means of addressing the rapid spread of the epidemic. Marks, writing in 2002, builds on these ideas and describes the situation in South Africa in the early 1990s as an ‘Epidemic waiting to happen.’ [8]

John Hargrove reviewing the historical development of the HIV/AIDS epidemic identifies the origins of the societal fragmentation which allowed for the diseases rapid development as being much earlier than the Apartheid era. The pattern of mass migration of men from South Africa and the surrounding countries to work on the goldfields of the Witwatersrand dates back to the turn of the century. Similar patterns were established to provide labour for mines and plantations in the British colonial areas of Southern and Eastern Africa. He describes this fragmentation of families, with wives and girlfriends left behind, as ‘.. a catastrophe waiting for something like HIV to come along’ [38].

An important development of these ideas was presented in a paper from 1995 by Sweat and Denison [39] published roughly 10 years into the global HIV/AIDS epidemic. The authors reflected that at the time there was a focus on AIDS as a disease of the individual and hence interventions to address it were based on trying to change individual patterns of behaviour. Whilst they accepted the importance of individual behavioural choices as the ultimate factor which determines whether an individual will be exposed to the risk of infection, they also called for a greater understanding of how that risk behaviour is influenced by the structural and environmental conditions experienced by the individual. They developed a hierarchical model in which causal factors are grouped from highest to lowest level as super-structural, structural, environmental and individual. Higher level factors work to influence those at a lower level in the hierarchy and ultimately influence individual behaviour. Super-structural causal factors are
the overall structure of the society, the political system and the way in which political power is managed to govern society. In the South African context, important factors in relation to this would be the deeply entrenched wealth inequity, unequal access to resources and pervasive attitudes such as the stigmatisation of those infected with HIV. Structural factors are the practical outworking of these super-structural factors. Historically in South Africa, Apartheid era legislation such as the Group Areas act or provision of single sex rather than family accommodation for male workers would be examples. The environmental factors describe the ways in which the structural factors lead to particular living conditions and patterns of access to resources for individuals. For example, the living conditions of mine workers, lack of access to sexual health services, poor provision of education particularly in rural areas and the level of gender inequity and gender violence experienced by individuals all contributed. Finally the individual level factors would refer to the experience of that environment by individuals and the way in which that is worked out in their behaviour. Sweat and Denison explore how this construct can be used to develop multi-level interventions to address HIV and how addressing social norms has the potential to enhance the impact of behavioural interventions. A more detailed summary of these definitions based on their paper is shown in Table 1.

Fenton writing in 2004 argued for a shift from the focus at the time on individual behaviours associated with the risk of infection [40]. This paper proposed that poverty was an important underlying factor affecting risk and that strategies to reduce poverty and inequality needed to be included in efforts to combat HIV. Parkhurst writing in 2014 develops this theme [41]. He contrasts the relative success of biomedical interventions against HIV such as male circumcision, pre-exposure prophylaxis and early ART initiation with the relative lack of success of behavioural interventions. To explain this he emphasises that as human sexual behaviours are deeply embedded in the social, economic, legal and political structures in which people live, an understanding of the influence of these structures on behaviour needs to be clearly conceptualised. He describes a framework which divides these factors into either drivers of risk behaviour or environmental factors, which mediate behaviour and affect an individual’s risk of infection with HIV. The causal pathway diagram developed by Parkhurst, which is reproduced in Figure 3, has a similar framework to the hierarchical model of Sweat and Denison and demonstrates in the case of poverty how this distal condition is worked out in its effects on the lives of individuals, leading ultimately to their behavioural choices.
Table 1 Definitions of hierarchical levels of causality in relation to HIV/AIDS in South Africa from Sweat and Denison (1995) [39].

<table>
<thead>
<tr>
<th>Causal Level</th>
<th>Definition</th>
<th>South African Examples</th>
<th>Change Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Super-Structural</td>
<td>Macro-social and political arrangements, resources and power differences that result in unequal advantages</td>
<td>• Economic underdevelopment&lt;br&gt;• Declining agricultural economy&lt;br&gt;• Poverty&lt;br&gt;• Wealth inequity&lt;br&gt;• Gender inequity&lt;br&gt;• Racial inequity&lt;br&gt;• Homophobia&lt;br&gt;• Stigmatisation of infected</td>
<td>• National and international social movements&lt;br&gt;• Revolution&lt;br&gt;• Land Redistribution&lt;br&gt;• War&lt;br&gt;• Empowerment of disenfranchised populations</td>
</tr>
<tr>
<td>Structural</td>
<td>Laws, policies, standard operating procedures</td>
<td>• Lack of family housing at mines&lt;br&gt;• Apartheid legislation&lt;br&gt;• Lack of human rights laws</td>
<td>• Legislative lobbying&lt;br&gt;• Civil and human rights activism&lt;br&gt;• Boycotts&lt;br&gt;• Legal reform&lt;br&gt;• Civil society activism</td>
</tr>
<tr>
<td>Environmental</td>
<td>Living conditions Resources and opportunities Recognition of individual structural and super-structural factors</td>
<td>• Hostels/single sex accommodation&lt;br&gt;• Lack of condoms&lt;br&gt;• Lack of sexual health services&lt;br&gt;• Living far from family/social support&lt;br&gt;• Gender inequity and gender violence&lt;br&gt;• Poor education provision in rural areas</td>
<td>• Community organisation&lt;br&gt;• Provision of accessible social/medical services&lt;br&gt;• Legal challenges&lt;br&gt;• Enforcement of laws&lt;br&gt;• Provision of more accessible health services</td>
</tr>
<tr>
<td>Individual</td>
<td>How the environment is experienced and acted on by individuals</td>
<td>• Social isolation&lt;br&gt;• Limited health related knowledge&lt;br&gt;• High risk behaviour&lt;br&gt;• Lack of perceived ability to change individual circumstances&lt;br&gt;• Age-differential sexual relationships</td>
<td>• Education&lt;br&gt;• Poverty alleviation programs&lt;br&gt;• Social grants&lt;br&gt;• Rewards/Punishments</td>
</tr>
</tbody>
</table>
Migration has had a profound effect on the social structure of South Africa and a significant impact on the development of the HIV/AIDS epidemic. It was estimated in 2006 that there were at least 2.5 million legal migrants in South Africa and a significant, though difficult to estimate, number of illegal or undocumented migrants [20]. One of the consequences of the Group Areas Act and other Apartheid era legislation was to force the black population living in rural areas to relocate, often to areas with a higher population density on more marginal land. This process, referred to as ‘villagisation’ forced a shift from a more agricultural self-sufficient way of life to a cash based rural economy. The need for cash and the lack of employment opportunities in rural communities was a trigger for the development of patterns of labour migration. To quote Shula Marks writing in 2002:

“The system of migrant labour and African rightlessness in so-called “white South Africa” was further entrenched to affect not only the mining industry, but also all
industries within “white South Africa”. Moreover, from the late 1950s, in pursuit of the government’s policies of Apartheid, South Africa was also the site of some of the most massive population movements in peacetime, outside of the Soviet Union. Over three million people were uprooted and their communities destroyed under a variety of Apartheid laws.” [8]

An early study recognizing the link between HIV and migration was carried out in a rural area of KwaZulu-Natal in 1992 [19]. This showed that the risk of HIV infection was 3 times higher for those who had recently migrated compared to longer established residents of the area. Similar associations have been reported in other studies in Sub-Saharan Africa [20].

The mechanism by which migration leads to an increased risk of infection was thought to be due to the mainly male migrants bringing the infection with them when they returned to their rural communities from the urban areas. It was known that the migrant men were likely to report a greater number of sexual partners thus exposing them to a raised risk of infection [42]. The relative emotional instability of the recent arrivals in the urban areas and their anonymity have both been put forward as reasons to explain this [20]. However studies by Lurie et al. in rural KwaZulu-Natal have revealed a more complex picture [42]. This work confirmed the higher HIV infection risk of migrants, with migrant men 2.4 times more likely to be infected than their non-migrant counterparts. It also identified that those who had made multiple moves were at a greater risk than those who had just migrated once. Presumably this is explained by the relative insecurity associated with regular relocations. However it was also found that for around one third of HIV discordant couples it was the female partner rather than the male who was HIV positive. This indicated that the females were also engaged in multiple concurrent relationships and that there is a need to understand the more complex sexual mixing patterns associated with this.

Coffee et al. carried out a modelling study in 2006 in order to better understand the interplay between migration and sexual behaviour in determining the development of the HIV epidemic [7]. This suggested that in the early stages of the epidemic, migration was an important driver for the establishment of the infection in rural communities. At later stages however it was changes in sexual risk behaviour that became more significant.

Another way in which migration had a profound effect on the consequences of the epidemic was demonstrated in a study by Clark et al. in the Agincourt HDSS study site in South Africa, the site of the present study [43]. It revealed that risk of dying from HIV related causes was significantly greater in the first 5 years of return to a rural community from an urban area than for those who had been settled for longer periods. This phenomenon referred to as ‘Returning home to die’ was thought to be due to the individuals returning to be cared for by their families when they became too ill to remain in work in the urban area. This provision of care placed an additional burden on rural households at the same time as they were affected by the economic impact of losing a bread winner [44].
Poverty and education

The relationship between wealth or poverty and the risk of becoming infected with HIV or dying of HIV/AIDS related illnesses is complex. One consensus that we can identify is that it is not wealth per se that is the main determinant of the relationship but rather the social, economic and environmental context in which the individual lives. This context may include many factors including the degree of wealth inequity in their home area, the stage of the epidemic and the availability and accessibility of healthcare providers offering HIV testing and treatment. Locally predominant beliefs regarding the mechanisms by which HIV is acquired and appropriate means of treatment are also important. Other factors such as gender related discrimination, power imbalances in sexual relationships and stigma experienced by marginalised groups also need to be considered in this context. A person’s level of wealth may have competing and contradictory effects at different stages of the pathway from HIV infection to death. It may be that a person of higher wealth has a higher risk of infection whilst at the same time having a greater ability to overcome the barriers associated with initiating and remaining on treatment. Another aspect is social capital, a term which refers to the number and strength of the social and familial relationships in which an individual exists. These networks can provide financial and emotional support at times when the person is unable to work due to the effect of illness or the need to attend healthcare appointments. They can also be an important channel by which individuals gain the information they need to make informed decisions on reducing the risk of infection and where to get tested and seek treatment if they do become infected.

Another important factor relates to how poverty is measured as the means of measurement may influence the associations seen. Howe et al. published an overview of different methods used to measure socio-economic position (SEP) in low and middle income countries [45]. The authors state the importance of understanding the nature of the particular measure used in an epidemiological study when interpreting the results found. In broad terms these can be divided into seven classes:

i) Asset based measures - here the household is ranked relative to others based on ownership of tangible assets, physical characteristics of the house such as the building materials used and the access household members have to services such as electricity, sewerage and water.

ii) Consumption expenditure - this measures how a household’s income is used to meet its material needs, typically through assessing what goods and services are purchased by the household. This is generally seen to have more value in low and middle income settings than purely looking at income coming into the household.

iii) Education - this is commonly used as a measure of SEP in both low and high income settings. In countries where education is not universally freely available, it provides a good proxy estimate for the amount of money a household has available to spend on school fees after their basic needs such as food, water and
shelter are met. This may be a limitation for its use in South Africa where education in the public sector is provided free up until the age of 18.

iv) Income - the cash flowing into a household. This is typically much harder to measure in low and middle income countries due to high levels of informal employment, seasonal employment and self-employment.

v) Occupation - this uses the nature of the work carried out by an individual as a measure of social standing. In some settings the divisions used may be very crude, such as employed vs. unemployed, and this can be problematic again due to the high levels of informal and self-employment.

vi) Participatory wealth ranking – in this approach community members, in consultation with one another, decide on the relative wealth rankings of different households in their community, an alternative method in which individuals give an assessment of how they perceive their social position in relation to others, may also be used.

Shelton et al. reviewed data from surveys in Kenya and Tanzania in 2005 [46]. In both cases there was a positive association between HIV prevalence and wealth. Similarly Mishra et al., reporting on a wider survey of data collected in Kenya, Ghana, Burkina Faso, Cameroon, Tanzania, Lesotho, Malawi and Uganda between 2003 and 2005, showed that those in the wealthiest SEP quintiles had the highest levels of HIV prevalence [47].

In 2008 Hargreaves et al. carried out a systematic review of studies investigating the relationship between HIV prevalence and educational attainment [48]. This review identified that the decreases in HIV prevalence over time were greatest amongst those who were the most highly educated. Developing these ideas, Hargreaves et al. analysed DHS survey data on HIV prevalence and educational attainment [49]. Based on this analysis they proposed that at earlier stages of the epidemic the wealthiest were at most risk whilst later, the burden fell most heavily on the poor. A possible explanation for the mechanisms underlying this could be that those of higher wealth are likely to participate in wider sexual networks and also be more likely to travel, perhaps between regions of lower and higher HIV prevalence. However these same factors and the economic buffer provided by wealth makes them more likely to be aware of and respond to opportunities for testing and treatment as they become available. They argue that this pattern would follow that set out by Victora et al. in the ‘Inverse Equity Hypothesis’ which proposed that new public health interventions are first taken up those who are wealthier and only later come to benefit the poor [50]. They suggest various methodological explanations for the different conclusions reached in previous analyses, for example; the geographical spread and sampling method used to identify the target population, the measure of wealth used and whether HIV prevalence or incidence is the outcome used. However in a more detailed later study that used data from 14 DHS surveys, Hargreaves et al. found inconsistent results in different settings [51]. Whilst there was some evidence for an association between changing HIV prevalence over time and education level, the nature of the change varied in different places. In some countries HIV prevalence was higher in the more educated whilst in others
HIV prevalence was greater among those with less education. Where a decrease in prevalence was seen, it did not only occur in the most highly educated. One difficulty in interpreting these findings is the use of HIV prevalence as a marker for changing risk associated behaviour. HIV incidence, the rate of infection, would be a better measure of the behaviour-poverty-HIV interaction but such data from longitudinal cohorts is relatively rare.

Parkhurst writing in 2010 argued that it is not wealth or poverty *per se* that defines risk status but rather the way in which wealth is associated with protective or risky behaviours, such as condom use and having multiple concurrent partners [52]. Fox carried out a series of cross-sectional analyses over many countries and regions in Sub-Saharan Africa [53]. The paper concluded that it was wealth inequality rather than absolute wealth that was the major determinant of the risk of infection. There was evidence that in poorer areas the wealthiest were at most risk whilst in relatively wealthy areas those who were poorer were at a higher risk. Connecting with the ideas of Lurie et al. [42], the paper further proposes that labour migration between the relatively poor rural communities and the wealthier urban areas may be an important means by which this inequality arises. A limitation of this study acknowledged by the author is that the data are cross-sectional and hence cannot account for time trends within regions.

Hargrove argues that in using data from multiple countries it is important to control for the stage of the epidemic within each country. He compared the Gini coefficient [54] (a measure of wealth inequality within a country) with the peak level of HIV prevalence for 28 Sub-Saharan African countries [38]. In bivariate analyses the Gini coefficient was associated with peak HIV prevalence but this association was lost in multivariate models. The percentage of males circumcised and the percentage Christian were associated with peak HIV prevalence. He concluded that

“It is thus hard to argue convincingly on currently available evidence for a causal link between a country’s HIV prevalence and the degree of income inequality.” [38]

One issue to consider in interpreting these studies is that the total HIV prevalent population is comprised of all those entering the population through infection minus all those leaving through death. If those of higher wealth are indeed more likely to gain access to treatment and are hence more likely to survive, one would expect the relationship between higher wealth and higher HIV prevalence to be maintained. Another problem particularly associated with cross-sectional studies is that one cannot elucidate the pathway of causality i.e. has poverty led to HIV infection, through influencing individual behaviour, or have the illnesses associated with becoming HIV infected led to poverty, through the economic impact of the disease on individuals and their families.

Bärnighausen et al. analysed data from a longitudinal population-based cohort in rural KwaZulu Natal to study the association between various measures of wealth or SEP and the risk of HIV infection [55]. Using data collected between 2004 and 2005 they found that each additional year of education lowered the risk of HIV infection. When they assessed the
relationship between infection risk and wealth it was found that those in the middle band of household wealth had a higher risk of infection than those in both the poorest and wealthiest strata. However no associations were found between measures of daily household expenditure, which is a measure of cash wealth as opposed to asset ownership, and the risk of infection. Their conclusion was that efforts to increase educational attainment were likely to be more effective than poverty reduction programmes in lowering the risk of HIV infection. They detected no gender differential in these relationships although they point out that their study may have been under-powered to detect such an effect. A much larger recent analysis of data from Botswana shows that each additional year of education gained by pupils, due to a change in the national education policy, led to an overall 8.1% decrease in the cumulative risk of HIV infection. A decrease in the risk of 11.6% was seen for females [56]. Cooke et al. studying the same population in KwaZulu-Natal as Bärnighausen et al. [55] carried out an analysis of the factors associated with the uptake of ART between August 2004 and December 2008 [57]. There was a small but significant difference in the percentage of males (19.2%) and females (21.8%) accessing treatment. There was no association found in a multivariate analysis with the levels of education or household assets. There was however a strong negative association with the distance individuals lived from a primary healthcare facility. As the Bärnighausen study covered only the last two years of the five year period used for the Cooke et al. study it is possible that the association between education and ART uptake may have changed over time.

Wringe et al. examined data on socio-demographic factors associated with ART uptake and coverage within four community HIV cohort studies in Southern and Eastern Africa [58]. Results showed that the relationship between level of educational attainment and ART coverage differed between the sites. In Karonga (Malawi) and Kisesa (Tanzania) there was evidence that ART coverage increased with increasing education although in the latter site the sample size was small. In Masaka (Tanzania) and Manicaland (Zimbabwe) there was some evidence of a decrease in ART coverage with increased education, but this was not a significant trend. In the same study a logistic regression model examined the likelihood of an individual having been screened for ART at two years after diagnosis of HIV. Only in Karonga were any significant associations with education level found and these indicated that those with five to eight years of education had the highest level of screening, eight years would indicate the completion of primary education. Those with either less than five years of primary education or at least some secondary education had lower levels of screening.

Tromp et al. reviewed six studies which assessed the relationship between different dimensions of SEP and utilisation of ART [59]. In two of these it was found that those with lower levels of educational attainment had less access to treatment. For the other four, which the reviewers deemed to be of higher quality, no relationship was found. Again some of the differences may be explained by the context and time periods in which the different studies reviewed were carried out and whether the measure of ART utilisation used was uptake or adherence. One of the studies reviewing data from a mobile HIV testing clinic in the Western Cape showed...
evidence that those who were employed had a lower likelihood of being linked to care, perhaps due to the difficulty of taking time out from work to attend a clinic. Peltzer and Pengpid carried out a review of studies assessing the relationship between HIV adherence and socioeconomic factors in low and middle income countries [60]. They identified 62 studies addressing various components of SEP: education, income, occupation or employment. Again different relationships were seen in different contexts, the results are summarised in Table 2. As can be seen no association was most common for each category, and a positive association was more common than a negative one. They noted that there was a high degree of heterogeneity between the studies and also that it had been impossible for them to carry out a formal systematic review due to the differences in how adherence was measured and the components of SEP estimated between the studies. Their overall conclusion, as with other investigators, was that the relationship depends on the cultural, geographic and economic contexts in which the study took place.

Table 2 Summary of studies on the association between the main components of socioeconomic status and adherence to antiretroviral therapy using data from Peltzer and Pengpid (2013)[60]

<table>
<thead>
<tr>
<th>Nature of association with ART adherence</th>
<th>Education N (%)</th>
<th>Income N (%)</th>
<th>Occupation N (%)</th>
<th>Socioeconomic status N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>10 (20.4)</td>
<td>15 (41.7)</td>
<td>3 (11.1)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>35 (71.4)</td>
<td>20 (55.6)</td>
<td>22 (81.5)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>4 (8.2)</td>
<td>1 (2.8)</td>
<td>2 (7.4)</td>
<td>2 (100.0)</td>
</tr>
<tr>
<td>Number of studies</td>
<td>49</td>
<td>36</td>
<td>27</td>
<td>2</td>
</tr>
</tbody>
</table>

An important factor to note in the context of South Africa that may lead to differences between associations seen there and those elsewhere is the wide availability of social grants. There are a variety of these: child support grants, disability grants and pensions being most common, particularly in rural areas. These grants are effective in lifting much of the population out of absolute levels of poverty and thus may alleviate some of the effects of poverty on HIV [61]. Another aspect of the relationship between poverty, social grants and HIV is highlighted by LeClerc-Madlala et al. [62] and Hardy and Richter [63]. In 2005 those who were HIV positive with a CD4 count of less than 200 cells/μl or who were seen from a medical assessment to have clinical signs associated with Stage 4 of the illness qualified for a disability grant [64, 65]. Thus for the poor there was a perverse incentive to default on treatment in order to drive down their CD4 count and retain access to the grant. The monthly sum of R780 (approx 110 US dollars at the time) meant that the loss of this grant would have a considerable adverse effect on the individuals receiving it and their dependants. The qualification criteria was subsequently amended in 2008 disallowing removal based on CD4 count criteria [66].
Gender and age

The associations between gender and the risk of HIV infection vary in different parts of the world. In higher income countries such as those in Western, Central and Eastern Europe, prevalence is higher amongst men than women whereas in low and middle income countries the situation is reversed. Across all low and middle income countries it is estimated that 52% of those living with HIV are female, whilst in Sub-Saharan African countries the figure is estimated to be 57% [67]. Worldwide amongst young people aged 15 -24 years, 0.6 % of males and 0.3% of females are estimated to be infected with HIV, whilst in Sub-Saharan Africa 3.1 % of females and 1.3% of males are HIV infected [68]. The explanations for this are in part biological and in part related to the social circumstances experienced by women.

In many parts of the world, particularly in low and middle income countries, there are major inequities between men and women in terms of employment, educational opportunities and power imbalances in relationships. The latter, caused by societal gender norms related to masculinity and femininity, have an impact on the health of both men and women. In a South African setting this often leads to male physical and psychological violence directed toward women [69, 70]. In addition the higher level of risk-taking behaviour amongst men, such as being involved in multiple concurrent relationships and unprotected sex increases the risk for their female partners. A recent systematic review has shown that the risk of acquiring HIV infection per sex act for a female is twice that for a male [71].

MacPhail et al. in a 2002 study in South Africa showed that age disparity in relationships, where younger women are in partnership with older men, is the single most important social factor predictive of the risk of HIV infection [72]. Maughan-Brown et al. using nationally representative data from South Africa looked at both age disparity and partner concurrency and this showed that younger women were more likely to be in relationships with older men and these men often had other concurrent sexual partners thus increasing the young women’s risk of HIV infection by linking them into sexual networks involving older people [73]. A common belief is that this is purely based on the cash and other material benefits young women receive from these older men, the so-called ‘Sugar Daddy’ theory. However studies by Ott et al. in a rural community in KwaZulu-Natal in South Africa suggest that the reasons are much more complex than this with age gaps more common in spousal relationships than casual relationships [74].

The relationship between intimate partner violence (IPV) and the risk of HIV infection has been demonstrated from an analysis of data collected from a five year longitudinal study in the Eastern Cape in South Africa, which showed that higher levels of IPV and a greater degree of power inequity in relationships were associated with increased levels of HIV incidence [75]. In contrast to this increased risk of HIV infection, women are more likely to know their HIV status, in part due to widespread coverage of antenatal HIV testing, and also more likely to be on treatment for HIV. A review of studies on the utilisation of HIV testing and counselling by Obermeyer et al. in 2007 showed a general pattern of low levels of HIV testing even in high prevalence countries and noted that relatively few studies had looked at the associations

25
between gender and behaviour related to testing [76]. They report that attitudes towards HIV risk are strongly influenced by gender, with men more likely to underestimate their risk of infection.

Lessels et al. carried out a survey in a rural area of KwaZulu-Natal in South Africa of the level of retention in HIV care for individuals who were infected with HIV but had not yet reached the threshold for ART initiation [77]. It was found that males were less likely to be retained in care than females. Cornell et al. reported on gender differences in survival among adult patients from a multicentre cohort study [78]. In this study, at treatment initiation, men had lower CD4 counts, were more likely to be classified clinically as WHO stage III/IV and had higher mortality. Additionally males were less likely than females to access ART and subsequently had higher levels of mortality when on treatment. Late presentation only partially explained the mortality differences. There was no evidence of gender differences in adherence as evidenced by comparisons of the level of virological suppression, although other studies have detected such differences. In order to better understand the results, they reviewed the gender specific mortality rates of HIV negative adults of comparable age to their cohort and found that males in this group had twice the mortality risk of females, due in particular to the occurrence traumatic deaths and non-HIV related tuberculosis. This may in part explain the differences in HIV mortality rates by gender seen amongst those how were HIV positive. A systematic review carried out by Druyts et al. found similarly that there was a significant increase in the risk of death for men when compared to women [79].

**Physical barriers to accessing HIV services**

A number of studies have investigated the relationship between physical barriers to accessing treatment and HIV/AIDS. Cooke et al. analysed patterns of ART uptake in a rural area of KwaZulu Natal [57]. They investigated various socio-geographic factors to see if they were associated with ART uptake in HIV positive populations. They found that the distance that individuals lived from the nearest clinic providing treatment remained strongly predictive of adherence after adjusting for other variables. The likelihood of taking up treatment for those 5 km away from a clinic was half that of those who lived close by. The authors concluded that these results pointed to a need for mobile and home based provision of care. Zachariah et al. carried out a study of patients with tuberculosis who were offered ART in a district hospital in Malawi [80]. They found that the likelihood of ART acceptance was linearly related to the transport cost to the hospital with patients who paid less than 50 Malawian Kwacha (MWK), approximately 0.5 USD at the time, for a one way trip to the hospital being four times more likely to take up ART than those paying more than 100 MWK. They concluded that decentralised ART provision would lead to a significant increase in take-up.
Sartorius et al. analysed data collected between 1992 and 2008 in the Agincourt HDSS study site in North-East South Africa, the site of the current study. They found that the straight-line distance that individuals lived from the nearest health facility was not a significant risk factor for all-cause mortality [81]. However at that time ART was not easily available from the local clinics, hence making it a measure not particularly representative of the extent of the barriers to accessing ART. Also the straight line distance measure used in this study does not reflect the true distance of travel to the clinics or include other barriers to access, such dirt roads or rivers that might be encountered.

In 2008 Moshabela et al. carried out a study of the effect on accessibility to healthcare of down-referral from hospital facilities to local clinics for HIV care. This was carried out in Bushbuckridge, an area that includes the site of the present study [82]. They compared down-referred patients treated in primary healthcare centres (PHC) with those treated at hospital. Using the healthcare accessibility framework developed by McIntyre et al. [83], they explored the influence of down-referral on availability, affordability and acceptability of care. The general conclusion was that down-referral had lowered access barriers with the cost of travel and the time associated with seeking healthcare reduced for down-referred patients. In a multivariate analysis there was evidence that long term adherence to ART was improved for the down-referred patients. Kunihira et al. carried out a series of qualitative interviews with key informants, healthcare workers, administrators, representatives of NGO and community organisations and individuals on treatment for HIV/AIDS [84]. The long distance of travel to treatment centres was identified as a key barrier to accessing ART by around half of those interviewed.

The impact of health related belief systems on the HIV/AIDS epidemic in Sub-Saharan Africa

In understanding the barriers to an effective response to HIV and AIDS in Sub-Saharan Africa it is important to understand the commonly held belief systems of the local populations and the way in which these belief systems have led to their understanding of HIV/AIDS and their response to it.

An ethnographic study was carried out in the site of the current study between 2002 and 2003, before ART was easily available in the public health sector [85, 86]. One aim was to understand how people understood and responded to illness. The investigators found that locally held folk beliefs were central to the process of health-seeking decision making. At that time there was little information available locally about HIV and AIDS so a framework of understanding involving ancestors, bewitchment and the violation of social and behavioural taboos developed. They found that choices about treatment were based on local knowledge of the relative efficacy of the available options for the particular medical condition. These options included biomedical care provided by the local hospitals and health centres, the use of traditional healers (both herbalists and spiritual healers) and Christian spiritual healing mainly through the dominant African Pentecostal churches in the area. The information used as a basis
for these decisions came through informal networks within the community, those individuals regarded as respected sources of information being referred to by the researchers as ‘therapy referees’ who were members of the social and kin networks of the individual. They identified that witchcraft was used as an explanation for otherwise inexplicable, mysterious and frightening occurrences. They contrast the situation in regard to TB treatment at the time to that for those with HIV/AIDS. In the case of TB a short course regimen had been available locally for many years and its effectiveness in curing those with TB was well established. Hence those with symptoms of TB were likely to go to the local health centres and hospitals for treatment. In contrast no effective treatment for HIV/AIDS was available locally at that time and so individuals looked to alternative treatment providers.

Fottrell et al. carried out an analysis of data reported in verbal autopsy interviews in the Agincourt HDSS study site between 1992 and 2006 [87]. Over the whole period, the term ‘bewitchment’ or associated phrases were found in the narrative description for 47.9 % of HIV attributed deaths. Comparing data from 2001 to 2006 with that from the earlier periods they identified an overall decrease in the number of deaths for which ‘bewitchment’ was attributed as a cause or explanation. There was no evidence that the decrease was greater for deaths assigned to be caused by HIV/AIDS than those attributed to other causes. Another factor that may explain the decrease in reports of bewitchment is that after democracy came to South Africa in 1994, there was an increase in politically motivated, violent witch hunts, leading to an upsurge in people’s consciousness of bewitchment. Over time and in part due to ongoing efforts toward post-Apartheid reconciliation there has been a reported decrease in witchcraft-related violence.

Niehaus, who has carried out long term ethnographic research in an area close to the Agincourt HDSS study site argues that in the pre-ART era, when being diagnosed with HIV was seen as a death sentence, then an explanation of the disease being due to bewitchment was easier to comprehend than other possible explanations, as it fell within the realm of pre-existing belief systems [88]. He also reported that a belief in bewitchment being causative didn’t necessarily preclude people accessing ART. Stadler, who carried out his fieldwork in an area close to that used by Niehaus, developed this idea. He described that, amongst his informants, witchcraft was not seen as causing AIDS but that it may explain why some were afflicted and others not [89, 90]. A commonly held view was that the stigma was based on judgment of the ‘immoral’ sexual promiscuity of the individuals infected. Niehaus presents an alternative perspective from research in the period prior to the availability of ART [91]. An HIV diagnosis was seen as an effective death sentence and placed the individual in a status referred to as ‘death before dying’ or ‘social death’. People known to be HIV positive were thought of as living corpses and thus shunned by their peers. Niehaus also reflected on the fact that a lack of treatment literacy was often identified as an explanation for the inadequate uptake and ineffective use of ART [88]. This led to a focus on education to improve people’s knowledge of the disease. He identifies other significant barriers which may be of greater importance, for example: a lack of political authorisation of the use of ART, social stigma, medical pluralism due to the availability of alternative plausible diagnoses, issues related to gender inequity and poverty. He also identified
the inadequate treatment available in under-resourced, overcrowded medical facilities as a major problem.

Liddell et al. carried out a study, in a rural area of KwaZulu Natal, of the association between belief in both ancestral protection from harm and traditional beliefs about illness with protective sexual behaviour [92]. They identified that older informants who held strongly to traditional beliefs were more likely to take precautions against AIDS, such as using condoms, perhaps due to an understanding that they provided physical protection against unclean elements entering their bodies during sexual intercourse.

**The history of demographic surveillance**

A major challenge in assessing the changing healthcare requirements of those living in low and middle income countries is the lack of population based health data from which an assessment of healthcare priorities can be made [93]. Ideally these countries would have universal population registries of the type seen in high income areas but these are often lacking or incomplete [94]. Solving this is a major challenge due to the logistical difficulties associated with establishing and maintaining such systems [95-97]. Whilst the goal of universal and accurate vital registration systems remains, in the mid-term at least other interim solutions are needed to provide the data required [98].

Household surveys such as those within the Demographic and Health Surveys (DHS) programme are available in some areas but coverage is inconsistent [99]. This makes it difficult to meaningfully and systematically monitor health metrics and assess whether global efforts, such as those toward the achievement of the Millennium Development Goals (MDG), are achieving their objectives [100]. Health and demographic surveillance systems (HDSS) have been established to address this data gap. HDSS survey sites provide complete coverage of all the vital events occurring in a defined population and in addition follow these events over an extended period [97]. The concept was pioneered at sites such as that at Matlab, run by the International Centre for Diarrhoeal Disease Research, Bangladesh, which has been carrying out continuous population surveillance in a rural area of that country since the early 1960s [101].

The initiation of an HDSS site will usually involve a baseline census including enumeration of all individuals and social structures resident in a defined geographical area. Typically an HDSS will also record familial relationships and marital unions. The required sample size will be driven typically by the need to ensure rare mortality outcomes are identified [97]. Having a study population of this size will also enable smaller sample populations to be selected from the overall population to enable the analysis of morbidity, socio-economic status and other determinants in the population.

At each household visit all entries into the population cohort, through birth or in-migration, and exits, through death or out-migration, are recorded. Sites will also often track all internal movements within the site enabling the HDSS to be used as a platform for analyses of factors
affecting individuals and how they change over time. Figure 4 presents gives a schematic overview of this process.

![Figure 4 Schematic describing the processes by which individuals enter and leave a Health and Demographic Surveillance System - adapted from Sankoh and Byass (2012) [97]](image)

The regularity of the household visits varies between different sites. The decision on frequency will typically involve a trade-off between resource constraints and scientific considerations [102]. A significant issue for sites with less regular visits is the risk of missing key events which may occur between visits. Identifying early neonatal deaths or miscarriages may be a particular problem for sites with less regular visit schedules as the entire course of the pregnancy and the early life of the child may fall between visits. The child may not be considered to be ‘fully born’ socially and hence a member of the household, until a certain period after the birth has elapsed. Thus the family would not report that child as a household member. To address this some HDSS sites record all known household pregnancies at each visit and follow up the outcome of these at a subsequent visit.

Another challenging issue is the need to assign accurately the length of time individuals actually reside in the HDSS. In some areas the majority of men and many women of working age may be employed away from the study area and yet still retain membership of the home household and return there regularly. This migration pattern is referred to as temporary, circular or oscillating migration by different authors. Ideally one would wish to record the actual time of residence at the study site in order to have accurate person time denominators for the calculation of birth, death rates etc. This can however be very difficult to achieve [103].
A key issue to consider in interpreting the results from an HDSS site is that of how representative the population of the site is of the surrounding populations about which one is seeking to make inferences. When considering this one must look at the demographic, socio-economic and health profile of the original population and how that population has been affected over time by the presence of an HDSS site. A difficulty in this is that often there are no data to assess these factors in other populations or where the data exists the quality may be poor.

As a result of a workshop held in 1998 in Dar es Salaam in Tanzania, an international network of HDSS sites, INDEPTH (International Network of field sites with continuous Demographic Evaluation of Populations and Their Health), was established. The network brought together the activities of a group of independent health research centres running HDSS sites. Currently the network includes 45 centres running 52 HDSS field sites in 20 low and middle income countries in Africa, Asia and Oceania [97, 104]. INDEPTH seeks to build capacity at the individual sites through technical training, scientific engagement and the establishment of cross-site research groups. Due to the independent development of these individual sites, data is collected and managed in a wide variety of data systems making multi-site data comparisons very difficult. To address this the INDEPTH iShare initiative has been set up to collate data from in a consistent format, provide rigorous data quality checks and create a data repository from which data extractions can be made to support multi-site analyses [97, 104, 105]. The repository includes both core demographic data and mortality related data from verbal autopsy surveys.

An interesting development proposed by Clark et al. is to extend the reach of an HDSS site over a larger population using statistical methods to sample people living in the wider area in a representative manner [106]. This has the potential to improve the generalisability of HDSS data.

Central to any HDSS study site is the data management system used to capture, clean and store the information collected at each observation. This in turn provides a platform to present the data for subsequent analysis. A major challenge in the development of such systems is that the data is longitudinal rather than cross-sectional and one must capture the dimension of time associated with the events recorded.

The original Matlab data management system was developed over 30 years using the best technology available at the time [107]. In this system extensive coding was required to ensure the integrity and consistency of the data captured. With the development of relational database management software (RDBMS), it became possible to produce a simpler data system in which many of the data integrity constraints are built into the design of the database. The Household Registration System (HRS) which is an example of this has been described by Benzler et al. [108]. This was adopted by many of the early HDSS sites. Central to the HRS database schema is a table describing the individuals. The data model then contains entities which describe the events experienced by the individuals and the episodes (periods of time) book-ended by these events. Episodes will include periods of residence in a location, periods of membership in a
social unit (family), marital relationships and pregnancies. The events recorded include births, deaths, in and out migrations. All visits are captured as observations which are linked to the data collected at those visits. Additional tables capturing information about the individuals would be built around this core structure. Within an RDBMS, relationships are defined between tables using common identifiers to link them. Data constraints such as those described above are programmed into the database and may be based on data in single or multiple linked tables. A schematic diagram of the data structure used in the Agincourt HDSS database which uses this architecture is shown in Figure 5.

![Diagram of the HDSS data model as used in the Agincourt HDSS](image)

**Figure 5 Schematic overview of the HDSS data model as used in the Agincourt HDSS**

An alternative temporal database design which has a simple underlying structure and is extensible allowing capture of different types of information in the same underlying data structure has been described by Clark [109, 110].

A high proportion of the effort in an HDSS data capture operation involves the maximisation of the quality of the data. There are two main dimensions to this; firstly ensuring that the data captured are exactly what was recorded from the household informant and secondly ensuring that these data are consistent with predefined logical rules and data previously recorded about that individual. The precise details will vary between sites but ideally will involve data checks
with feedback loops to address data quality issues at all steps in the data collection and capture process, this would typically include:

- Confirmation of key data from current and past visits (e.g. dates of birth and household memberships) during the household interview
- Checking of forms in the field by a data supervisor
- Constraints at data entry to prevent the capture of inconsistent data
- Running of queries against the database subsequent to data entry

At each stage a set of procedures including checks of the forms for transcription errors and return visits to the households may be triggered to resolve data errors and inconsistencies.

**Verbal autopsy**

For 75% of the world’s population no systematic mortality surveillance is available [111]. Verbal Autopsy (VA) is a process to ascertain a deceased’s cause of death where a post-mortem examination has not been carried out. In a VA enquiries are made about the circumstances surrounding the death, after which the most likely cause is derived.

The historical development of the method can be traced back to 17th Century London where death searchers visited the homes of the deceased to make enquiries about the circumstances surrounding the death. Their emphasis was on identifying deaths due to communicable diseases. In many parts of the developed world systems of universal death registration have replaced this, whilst in less developed areas similar processes are still used [112].

The modern development of VA methods began in Africa and Asia in the 1960s. The process involved trained clinicians carrying out interviews at the homes of the deceased. Over time this process began to be replaced by questionnaire based VAs, which allowed for the systematic recording of signs, symptoms and circumstances associated with the death [102, 112]. With this approach the procedure was more standardised and reproducible and could be carried out by non-clinically trained fieldworkers, allowing for many more deaths to be recorded with the same resources. The development of the VA instrument came from the needs of individual research projects at different sites and led to many different highly diverse questionnaires being developed. In 2006 a survey of sites applying VA methods identified 18 different questionnaires in common use [113]. Efforts have been made through bodies such as the World Health Organization (WHO) and INDEPTH to develop standardised questionnaires. In 2012 following extensive consultation and a review of questionnaires in common use, the WHO released a VA tool with the aim of wide implementation and harmonisation of methods. The tool included three different VA questionnaires for use with different age groups, an open narrative section and questions designed to elucidate the presence or absence of particular signs and symptoms in the deceased’s final illness. In addition questions related to lifestyle and disease associated risk factors were also included [114]. A revised version of the instrument was released in 2014.
It has been recognised that it would be useful to record additional information about the circumstances associated with the death of an individual in order to understand the impact of other non-biological factors. Examples of such additional information could include any barriers to receiving effective care experienced by the individual during their final illness. Gaining knowledge of the extent to which these factors are associated with death can empower communities and health providers to participate in interventions that are aimed at addressing potentially modifiable factors. To address this gap, ‘social autopsy’ methods have been developed and are being implemented in some settings [115]. Social autopsy has developed from models which seek to understand the care seeking behaviour of individuals. These models focussed on maternal and child health [116]. A group of sites within the INDEPTH network have created a social autopsy tool for use with deaths occurring in those aged 5 years and under [97]. This has been piloted and the results analysed in Ghana, Uganda and Guinea Bissau [117]. A circumstances of death section, which collects relevant social information, is included for all age groups in the revised WHO VA questionnaire [118].

Subsequent to the VA interview it is necessary to determine the cause of death. The most common method that has been used is for experienced clinicians to review the VA forms and assign what is, in their opinion, the most likely cause of death. Typically two clinicians assess the forms independently. Where the assigned causes differ, the original clinicians will discuss the case and aim to reach consensus. Typically an arbiter would be used where consensus cannot be achieved. This technique is referred to as physician coded verbal autopsy (PCVA). Following what is normal clinical practice, the PCVA approach will usually focus on the data collected in the narrative sections of the VA form. This has the advantage of a more nuanced approach to the diagnosis, but can lead to subjective conclusions as the diagnosis will be based on the experience of the individual clinicians and may be subject to change over time [111].

An alternative to PCVA which has the potential to overcome some of these limitations is to develop algorithms or probabilistic models which can be used to analyse the symptomatic data collected in the VA questionnaire and assign causes of death. These are referred to as computer coded verbal autopsy (CCVA). A variety of algorithms have been developed which can be divided into two types; those that seek to capture expert opinion and those that are data-derived. The significant difference between the two approaches is that in the former the relationships between symptoms and causes are derived from the expert opinion of clinicians, whilst the latter typically use data from hospital based verbal autopsies. In this case the assumption must be made that the hospital deaths are representative of deaths occurring in the community and hence that the distribution of symptoms and causes is representative of that in the wider population [111].

Amongst the data derived methods, three have gained prominence in recent years. The Tariff method uses a training set of cases with known causes derived from hospital autopsies [119]. The likelihood of an association of a particular symptom with a known cause is derived. These likelihoods or weights are summed to allow the calculation of the most likely cause. The Random Forest method uses a more complex machine learning approach to develop rules...
which use the signs and symptoms recorded to distinguish between all possible pairs of causes [120]. The results of these pair-wise comparisons are then compared to derive the most likely cause. The King-Lu method derives aggregate rather than individual data on the distribution of causes in a particular population [121]. One difficulty of the method is that to break down the CSMF by factors such as age and gender, it is necessary to train the model on a sample which has those particular characteristics.

The probabilistic methods in contrast seek to simultaneously calculate the probability of all possible causes based on the signs and symptoms identified and through this assign causes of death with an associated likelihood of their occurrence. The InterVA model uses this approach and has been successfully applied in a variety of settings in low and middle income countries [111]. The InterVA model is based on Bayes’ theorem which links the probability of an event occurrence given a particular circumstance with the unconditional probability of that event and the conditional probability of the circumstance given the event. In InterVA the event would be the particular cause of death and the circumstances would be the signs and symptoms occurring in the period leading up to the death [122]. In order to create the underlying probabilistic model a group of experienced clinicians developed a set of prior probabilities of the occurrence of particular symptoms for a range of causes of death. The latest version InterVA-4 was released in 2012 and is freely available for download from www.interva.net. The symptom categories used are based on those in the WHO 2012 VA standard [114].

For the rigorous validation of CCVA methods one would ideally need a gold standard data set for which all causes have been assigned, symptomatic data collected and which is representative of the deaths in the community where the comparison is to be made. A validation dataset has been derived by the Population Health Metrics Research Consortium (PHMRC) [123]. However possible limitations in the quality and representativeness of that dataset have been identified [124]. Murray et al. have used this dataset to compare the Tariff, King-Lu, Random Forest and InterVA algorithms. Their analysis indicated that the Tariff method performed as well or better than the other methods tested [123]. Another comparison of four different CCVA methods with PCVA used a validation dataset of 24,000 deaths occurring in low and middle income countries [125]. Again Random Forest and Tariff were compared with InterVA and King-Lu. In this study the performance of each method was similar. King-Lu gave the highest Cause Specific Mortality Fraction (CSMF) concordance of the methods studied, however that method cannot calculate individual cause probabilities. A study comparing the partial chance-corrected concordance of CCVA and PCVA found overall a very similar performance for Random Forest, Tariff and InterVA [123]. Other comparisons of PCVA with InterVA-4 have been made. The most recent and largest compared CSMF for PCVA and InterVA for 54,000 deaths from six settings in Asia and Africa [126]. Calculations were made of the concordance correlation coefficients (CCC) for the CSMF values calculated using the two methods. Overall the results showed a high degree of correlation, with a CCC of 0.831 for the pooled data set (a value of 1 would indicate a perfect correlation). However the site specific data indicated that there was considerable variation in the CCC value between sites. For the two South African sites the CCC values were 0.588 and 0.912 respectively. A lot
of this variation however could be explained by the approach taken by clinicians in assigning deaths to HIV and TB. This is particularly problematic in a high HIV prevalence context as other studies have shown the rate ratio of those HIV positive to HIV negative among those dying of pulmonary TB was 63.9 [127]. When HIV and TB deaths were combined into a single category the CCC values for the two sites were virtually identical.

An earlier study compared the proportions of deaths assigned to have been HIV-related each year by the physicians and InterVA. This analysis used the previous version of the InterVA model, InterVA-3[128]. The results showed a high degree of similarity between the two methods. The major difference however was that in the early years of the HIV epidemic a higher proportion of deaths were assigned to be HIV-related by the InterVA model than the clinicians. This is probably explained by the quite understandable limitation of physician assessment that clinicians are unlikely to assign as a cause a previously unknown or very rare condition. Thus with hindsight and the lack of subjectivity of the InterVA approach we can review data retrospectively and gain a clearer understanding of the origins and development of disease epidemics. Given the differences in approaches to diagnosis in which physicians use almost exclusively the open narrative in the VA to determine cause and the computer methods use closed questions the degree of correlation seen here is impressive.

One way to overcome the challenges of comparing the predictive methods would be the derivation of a community derived autopsy data set in which autopsies are carried out for deaths occurring in the community rather than in hospital. However the practical and cultural challenges in carrying out such a project have yet to be overcome.

In summary it is not possible from the evidence of comparative studies that have been carried out to formally decide which method is the ‘best’ or most accurate. The two methods developed by PHMRC appear to perform much better on the data set from which they were built than when applied to data derived from a variety of other sources.

Most verbal autopsy based investigations have been carried out in relatively small local study sites, and results from the isolated sites can then be extrapolated through modelling to provide national or regional data. An example of the use of VA on a broader basis linked to civil registration of deaths is the Indian million death study in which to date 200,000 deaths have been assessed by VA and physician assessment [129]. This shows the potential for the wider application of the technique. The VA instruments have also been implemented on mobile platforms such as tablets or mobile phones, which have the potential to lower the costs of applying the technique and hence achieve a wider coverage than is possible using paper based forms and manual data entry [130].
METHODS

Study setting and the establishment of the Agincourt HDSS study site

This study was located in a rural area of North-Eastern South Africa. The majority of the population of the study area are ethnically Tsonga or Shangaan, with ancestry in the Gaza province of Mozambique. During the 1820s there were significant population shifts in the region as the native Nguni groups were forced into modern day Mozambique by the invading Zulu armies. This movement led the Tsonga speaking peoples in turn to cross the Lebombo Mountains and settle in this region. Around 1900 there was a largely peaceful co-existence between the local tribal chieftains and the white Afrikaner farmers who had settled in the area. The Afrikaner population of the region were descendants of the “Voortrekkers”, who had migrated to flee the British rule in the Cape colony in the 1830s and 1840s. In the early part of the 20th Century the black population were forced off their traditional land as the farms expanded [6]. In 1898 the Sabi Game Reserve, which borders the site of the current study, was established in the area. This was later expanded to become the Kruger National Park which was created as a reserve to protect wildlife and now occupies an area of almost 20,000 km², acting as a buffer between the eastern border of South Africa and Mozambique. The creation of the Kruger National Park resulted in further relocations of both black ethnic groups and white farmers.

The Gazankulu homeland in which the study site was formerly located was assigned to the Tsonga speaking peoples. It was comprised of three separate pieces of land, two close to the Kruger National Park and another near Tzaneen with a capital in Giyani see the map in Figure 6. The goal of this process was to create a series of semi-autonomous territories, dependent economically on South Africa [6]. In the area comprising the Agincourt HDSS study site, former white owned farms were compulsorily purchased by the South African government and individual villages established with surrounding open land which could be used by community members for cattle grazing and firewood harvesting. In some cases there were existing villages which predated the forced resettlement and which formed the basis of the new communities [131].

During the mid 1970s politically active medical students began to make visits to rural hospitals in the homelands. At the University of the Witwatersrand these rural placements later became part of the formal curriculum for medical students. After recognising the need for the development of rural health systems, John Gear, the first Professor of community medicine at the University, established the Health Systems Development Unit (HSDU) based at Tintswalo hospital in Acornhoek in Gazankulu. Efforts in the mid 1980s to improve the quality of service in healthcare were only of limited success, due in part to under-resourcing and a resistance to change amongst those responsible [132]. In the early 1990s Stephen Tollman and Kathleen Kahn arrived at HSDU and established a new health systems development programme. The
approach taken was based on experience gained through visits to established Health and Demographic Surveillance Sites (HDSS) including Matlab in Bangladesh and Niakhar in Senegal.

They recognised that in order to understand the effect of health interventions at a population level it was necessary to establish a platform which could monitor demographic changes in the population. The HDSS model was ideal for this and as a result the Agincourt HDSS was established in 1992, with the baseline census to enumerate the population carried out between 1992 and 1993. As an example of the paucity of data for planning health research at this time, local development agencies had estimated the population of the Agincourt sub-district to be between 30,000 and 35,000, following the baseline census this was adjusted upwards to 58,000. The goals of the Agincourt HDSS were to develop a research site with the infrastructure to introduce and evaluate health programmes and to provide population based data to form the basis of such evaluations [133].

Figure 7 is a contemporary map of South Africa showing the new provisional boundaries introduced after 1994. Figure 8 is a detailed map of the Agincourt HDSS study site and the surrounding area.

**In-migration to the Agincourt region from Mozambique**

From the late 1880s until the mid 1980s many migrants moved into the area from Mozambique, mainly to seek work on farms and in the mines. These in-migrants formed links with existing Tsonga speakers who had settled in the area prior to the modern day borders between South Africa and Mozambique becoming established. There was a significant decrease in the numbers of Mozambicans employed on the mines in South Africa in the mid 1980s, and many of these unemployed miners who were often married to South African women settled in what was then Gazankulu. Subsequently during the decade-long civil war in the mid 1980s many refugees arrived in the area from Mozambique. Due to the long-standing kin relationships there were already links to the existing population of the area. A number of the villages assigned land for the establishment of refugee communities. These ‘Mozambican’ villages were and to an extent still are characterised as having higher levels of poverty and poorer infrastructure than the neighbouring South African communities [134]. Due to this continuous process of migration and assimilation into the host community, the distinction between Mozambicans and South Africans is not always clear with roughly half of the Mozambican headed households being located in a mixed neighbourhood with at least one adjacent property having a South African household head. In the distinct ‘Mozambican’ villages around 25 % of the households were in 2003 South African headed [134]. Through formal and informal processes since 1995 over 85 % of those arrived in South Africa after the civil war have attained South African citizenship and the right to access social grants and employment [135].
Figure 6 Map of South Africa showing the provincial boundaries pre 1994 and the designated homelands with the location of the Agincourt HDSS study site indicated.

Figure 7 Contemporary map of South Africa showing the new provincial boundaries.
Figure 8 Detailed map of the Agincourt HDSS study site and the surrounding area
Characteristics of the Agincourt HDSS study site

As of July 2011, the Agincourt HDSS study site consisted of 27 villages and occupied an area of roughly 420 km². The total population was approximately 90,000, of which about 67,000 (75%) were permanent residents, defined as spending more than six months of each year resident in the site. The residents lived in approximately 16,000 households [136], a household being defined not necessarily according to kin relationships, but rather as a group of individuals who eat together. As can be seen from Figure 9 a high percentage of adult males (41% aged 15-39 and 58% aged 40-59) and a somewhat smaller percentage of adult females (25% aged 15-39 and 23% aged 40-59) were in 2011 classified as not being permanent residents. The majority of these were labour migrants either employed or seeking employment away from home. Thus many adult males are not regularly resident in the household in which they reside in the study site. Life expectancy at birth decreased from 72.7 years for females and 68.2 years for males to 59.4 and 53.3 years respectively between 1994 and 2004 (Figure 10). This was probably largely due to the high level of HIV/AIDS related mortality prior to the ART becoming available in the area. By 2011 life expectancy had increased to 65.8 years for females and 59.4 years for males (Source - INDEPTH Stats data repository http://www.indepth-ishare.org/indepthstats/ [137]).

Electricity is now available in most villages although the percentage of households using it for heating and cooking varies considerably between villages. Water is generally supplied through public taps or stand-pipes from which it has to be transported to the dwelling, although some households do have their own connections. The roads are mainly gravel and subject to erosion, particularly in the rainy season, although some major roads across the site now have a tar surface. Each village has a primary school and most a high school [136]. The level of school attendance is high, however the overall the level of educational attainment is low. Bushbuckridge, the district in which the site is located, had the lowest matriculation (school leaving examination) pass rate in the whole of South Africa in 2010, with only 40.8% of learners successfully passing the exam [138].

There were in 2011 six primary care clinics and two health centres within the site and three district hospitals serving the site population and between 25 and 60 km away [136]. Local employment opportunities are very limited as evidenced by the high proportion of adult labour migrants. A variety of leadership structures operate in these communities. These include the Indunas or traditional chiefs and their representative councils, the elected ward councillors and the Community Development Forum (CDF). The latter is an unelected voluntary body which includes in its membership representatives of a variety of community organisations in each village. In general the CDF is consulted on issues of community development, particularly in terms of the provision of infrastructure such as water, electricity and sanitation. The traditional structures tend to be consulted with reference to land issues and social concerns [139].
Figure 9 Percentage of the total population of the Agincourt HDSS study site who were classified as permanent residents as of July 1st 2011

Figure 10 Changes in life expectancy at birth Agincourt HDSS study site (Source = Indepth Stats http://www.indepth-ishare.org/indepthstats/)
Spatial Analysis and its relevance to this study

In spatial epidemiological studies investigators seek to understand the associations of place with disease outcomes. Identification of such links can then be used to identify the factors associated with a particular place, which may be social or environmental, and may be causally associated with the measured outcomes.

John Snow, a medical doctor and anaesthetist working in London in the early to mid 1800s, was suspicious that contaminated water supplies were responsible for the transmission of cholera. This was contrary to the prevailing belief at the time that the cause was ‘miasma’, defined as – “a poisonous atmosphere formerly thought to rise from swamps and putrid matter and cause disease”\(^1\). During a dramatic and sudden cholera outbreak in Soho in 1854 he obtained details from the death records of the precise place of residence of each person who died of the disease. The majority of cases were in people living close to a particular water pump located on Broad Street. More detailed investigation revealed that some of those living closer to other pumps chose to use the Broad Street pump out of preference, as they preferred the taste of the water. Amongst the 535 inmates of the workhouse close to the pump there were only five deaths, far fewer than expected, but it turned out that the workhouse was supplied by its own well and did not use the pump, this was also the case for workers in the nearby brewery. Subsequently it was found that the source of water for the pump was close to a sewer which was the probable source of contamination. This investigation provided strong evidence for Snow’s belief that an infectious agent was responsible for the transmission of cholera [140, 141]. The evidence was used to convince the local parish’s ‘Board of Governors and Directors of the Poor’ to remove the pump handle to prevent more water being drawn. Snow was not the first to carry out a spatial epidemiological investigation, however his elegant and rigorous study is seen by many as the foundation of modern epidemiology and in particular spatial epidemiology [140].

The historical context of colonial history and Apartheid in South Africa have had a profound effect on where infrastructure and resources are located and the geographic space in which individuals reside [81]. This has led to differences in, amongst many other things, educational and employment opportunities and the quality of healthcare available. Household location also determines one’s access to land for agricultural purposes, natural resources and clean water. These factors, mediated by other individual and household level determinants, can affect a person’s health and wellbeing. Spatially associated differences in ecological and environmental characteristics and resources have been identified within the Agincourt HDSS study site [142].

Spatio-temporal clustering analysis using Kullendorff’s spatio-temporal scan statistic as implemented in the SaTScan\textsuperscript{TM} software v9.1 [143] allows one to identify whether morbidity or mortality patterns indicate that there are groups or clusters of cases which are more concentrated than would be expected by normal chance variation. This can be conceptualised

as taking a number of cross-sectional maps of an area on which all cases are identified and stacking them one above the other. Each map would represent a time range (e.g. year or month). A cylindrical scanning window which varies in diameter (representing the spatial extent of the cluster) and height (representing the time period over which the cluster occurred) is then moved across the stack of maps. Where a maximum likelihood ratio test shows the number of cases within the cylinder to be significantly greater than that outside, a spatio-temporal cluster is identified. A schematic representation of this is shown in Figure 11.

The technique has been used in a wide variety of applications. For example Cuadro et al. applied the technique to DHS data to identify spatial clustering of HIV cases across Sub-Saharan African countries [144]. In an exploratory analysis Sartorius et al. identified age-specific all-cause mortality clusters within the Agincourt research site. The statistically most significant clusters tended to be located in the south and east of the area [145].

![Figure 11 Schematic representation of Kulldorff's spatial scan statistic](image)

**Micro-epidemic analysis**

Another aspect of the investigation of spatial patterns of morbidity and mortality is the identification of local variations in patterns of disease, which may be masked if an analysis is only made at national, regional or district levels. A number of studies have highlighted the need to understand such micro-level variation in order to effectively target interventions [146-150]. Wand, Ramjee et al. [148, 149, 151] identified high prevalence spatial clusters of HIV and other sexually transmitted infections in the Greater Durban area of KwaZulu-Natal in South Africa and suggested that the identification of these be used as a basis for targeted biomedical, behavioural and structural interventions against HIV and STIs in these communities.

Tanser et al. [150] reviewed a number of recent studies investigating the occurrence of HIV sub-epidemics in generalised epidemics. In many of these studies it had not been possible to provide good estimates of HIV infection rates in hard to reach and typically high risk population groups such as men who have sex with men (MSM), people who inject drugs
(PWID) and Commercial Sex Workers (CSW). It is thought that groups such as these may act as drivers for the epidemic in the general population. They identified a number of studies in which high risk population groups were defined on the basis of geography. This again suggests that interventions targeted towards geographically defined high risk groups may be a particularly effective strategy to address the epidemic in the general population.

A report from UNAIDS also develops this theme [152]. The report recognizes the potential value of the use of data describing spatial patterns within an HIV epidemic for the identification of gaps in the provision of HIV services and shows, from a number of case studies, the value of understanding the geographical locations of cases associated with the epidemic at different levels of spatial resolution [152]. They note that there is often a spatial overlap between high levels of HIV prevalence and the presence of hard to reach groups or ‘key populations’. The report also emphasises that these potential benefits must be balanced against the need to protect the rights of individuals. These ethical issues are very important due to the risk of stigmatisation and potential harm for those individuals living in regions which may have a relatively high HIV prevalence compared to surrounding areas. This risk must however be set against the potential benefits if this data is used with appropriate safeguards in place.

The UNAIDS briefing document also reports on initial modelling studies on data from Kenya. These indicated that a package of specific targeted interventions including; early treatment, behaviour change, condom promotion, pre-exposure prophylaxis, medical male circumcision and programmes targeting key populations could avert an additional 600,000 HIV infections by 2030 for same budget currently dedicated to HIV prevention efforts [153].

In India a network of 1300 sentinel HIV surveillance sites has been developed allowing fine grained data on local epidemics to be collected. Through this it was identified that MSM in the south of the country and PWID in the north-east had a higher HIV prevalence than the general population. This enabled prevention efforts to be targeted to the most heavily affected regions [154]. Studies in urban areas of Nigeria [155], Uganda [152] and Kenya [156] identified particular areas of cities where people engaged in a variety of behaviour which increased their risk of HIV infection were found to congregate.
RESULTS

The four papers included in this dissertation describe the development of a local HIV epidemic and also explore changes in health seeking behaviour of the population over time. The results are described in detail in the original papers, what follows is a brief summary.

In interpreting the results a key contextual factor is the time at which ART became easily accessible in this community. Voluntary Counselling and Testing (VCT) for HIV became available in 1999 [157]. Two public hospitals whose catchments include the population of the study site began to provide ART in 2004, although initially the capacity was somewhat less than the demand and there were drug stock-outs and long waiting times for treatment initiation. Within the area covered by the Agincourt HDSS study site, the Bhubezi Community Health Centre (a public-private partnership) and three public sector health centres began to provide ART in 2007. By 2009-2010 patients could be initiated on treatment and return for follow-up monitoring at Bhubezi. This was also possible at the three public health centres, though resource constraints meant that the numbers for whom care could be provided at these facilities was limited. Also by this time many of the services provided in the public hospitals had improved considerably.

In the following sections the term ‘HIV/AIDS mortality’ is used to refer to all deaths for which the InterVA-4 model assigned either HIV/AIDS or pulmonary tuberculosis as the most likely cause. In some cases this terminology differs from that in the original published studies but is used here for consistency. This categorisation has been used in all the analyses in this study due to high level of HIV and TB co-morbidity found in this region [127]. Also the alphabetic codes used to identify villages differ between papers I and II to further preserve anonymity.

Paper I compares HIV prevalence for the Agincourt HDSS study site in 2011 with national data for 2012. Prevalence in Agincourt was higher than that nationally in all adult age groups, particularly so for those aged over 30. This study then provides an analysis of the trajectory of the HIV/AIDS epidemic for each of the villages as measured by changes in the HIV/AIDS mortality rates and compares these with the pattern for the whole site. The overall epidemic profile shows a small number of initial cases in 1992-1993 when the surveillance commenced reaching a peak rate around 2007. In general each individual community follows the pattern for the whole site with no evidence that the HIV/AIDS epidemic started in one area then spread from there to other regions of the site. Thus this can be described as an initially dispersed rather than concentrated epidemic. It is likely that the epidemic was introduced by returning male labour migrants who spent the majority of their time working mainly in the mines elsewhere in the country. There is some evidence that village S has followed a slightly different profile in recent years with no evidence for a decrease in the rates since 2007. For those aged less than 40, females had higher rates of HIV/AIDS mortality than males for the whole period of the study (1992 to 2013). In the older age groups (40 to 59 years and 60 years or greater) males had higher rates than females. This is despite the fact that at least in 2011 we see that the HIV prevalence for both groups was similar. There is also some evidence that at the epidemic’s peak those of Mozambican origin had higher HIV/AIDS mortality rates than their South
African contemporaries, though the differences are not statistically significant. By 2013 the rates for the two groups were very similar. It was also estimated that just over 60% of the total number of deaths occurring in the research site between 1992 and 2013 could be attributed directly or indirectly to HIV infection.

**Paper II** identified an overall decrease of 30% (90% CI 23% - 36%) in HIV/AIDS mortality amongst adults in the study population between 2007-2008 and 2009-2010. The time periods were chosen to represent periods just before and just after ART became more easily available. There were no significant changes in mortality rates associated with causes other than HIV/AIDS. The differences between cause of death categories and the fact that the decrease in HIV/AIDS mortality rates occurred at the same time as ART became available in the public health clinics in this area provides evidence that the ART roll-out rather than general improvements in the healthcare provision was the cause of the decrease in HIV/AIDS mortality. The range of mortality rate ratios varied from 1.35 (90% CI 0.78 - 2.34), a 35% increase, for village Q to 0.37 (90% CI 0.15 - 0.91), a 63% decrease, for village Y. This range demonstrates the high degree of heterogeneity between the villages. Notably two other villages H and L with rate ratios of 0.38 and 0.43 respectively are geographically close neighbours to village Q showing that this heterogeneity can be seen over short distances. These results are shown in detail in Table 3 which is from the paper’s online supplementary materials.

In the gridded analysis there was evidence for significant heterogeneity in outcomes over time. For example, the highest rate ratios for HIV/AIDS mortality between the latter and earlier periods were seen in two grid squares overlapping the Bhubezi Community Health Centre towards the east of the study site, whilst a pair of grid squares at the extreme east, 2 km away, showed significant decreases in the rate over the same period. This is shown graphically in Figure 12 reproduced from paper II.

These results provide evidence for large variations in HIV/AIDS mortality rates within a small geographic area. It can also be seen that factors other than the distance an individual dwelt from the nearest clinic or health centre were likely to be important determinants of the variation in rates. Hence other geographic and socio-demographic factors need to be identified to explain these variations.

**Paper III** identifies determinants associated with the risk of dying of HIV/AIDS between 2007 and 2010. It was found that there was a reduced risk of HIV/AIDS mortality for; those living less than 5km from the Bhubezi Community Health Centre, women, those living for more than six months of each year in the study site, members of households owning motorised transport and those living in households with higher levels of socio-economic position or education. In addition, the multivariate analysis found that there was an increase in the risk of dying of HIV/AIDS for those of South African origin compared with Mozambicans. No association was found between the taxi fare to reach Bhubezi from an individual’s home and the risk of dying of HIV/AIDS. A separate model was developed including the same determinants and with an outcome of mortality due to causes other than HIV and TB. The associations seen in this model were similar to the previous one. In order to assess whether any of the determinants had
significantly different associations in the two models a third model was developed in which the outcome was all cause mortality and the category of the cause of death (HIV/AIDS vs. other causes) was included as a determinant. This was introduced as an interaction term with each of the variables of interest from the other models. It was found that the rate ratio for males compared to females was 1.12 (95% CI = 0.99-1.27) times greater for those dying of other causes compared to those dying of HIV/AIDS. The rate ratio for those of South African origin compared to those of Mozambican origin was 0.85 (95% CI = 0.75-0.98) times lower for those dying of other causes compared to those dying of HIV/AIDS.

Paper IV investigates the changing patterns of health seeking behaviour over the periods before, during and after the roll-out of ART. For those dying of HIV/AIDS there was a decrease between 2003-2005 and 2009-2011 in the reported use of traditional or herbal remedies during the course of the final illness from 77.5% to 38.5%. For those dying of other causes the decrease was from 53.6 % to 23.6 %. In each of the three time periods into which the data was aggregated, those dying of HIV/AIDS were more likely to report any use of traditional or herbal remedies than those dying of other causes. In 2003-2005 the families of those who ultimately died of HIV/AIDS were more likely than those dying of other causes to report that traditional or herbal remedies were used as the first treatment option. By 2009-2011 it was more likely that those dying of causes other than HIV/AIDS were reported to have used traditional remedies first. The number of deaths where dual use of traditional and biomedical treatments was reported had decreased for both cause of death categories. However by 2009-2011 dual use was still reported for more than 30 % of those dying of HIV/AIDS. Figure 13, reproduced from paper IV, provides a summary of these results.

In a multivariate analysis using logistic regression, three determinants were found to have a statistically significant association with the likelihood of traditional or herbal medicine use during the final illness. These were an individual’s country of origin, with those of Mozambican origin more likely to report the use of traditional or herbal medicines than South Africans, the duration of the final illness, a longer duration made use of traditional remedies more likely and the calendar time period, use of traditional remedies decreased in later time periods.

There was evidence that the difference between those of Mozambican and South African origin decreased over time. There was no evidence that socioeconomic position, the level of educational attainment or the age of the deceased were associated with an individual’s likelihood of using traditional healthcare.
Figure 12 HIV/AIDS mortality rate ratios for 2009-2010 compared to 2007-2008 for the Agincourt HDSS study site from Paper II.

Figure 13 Changing patterns of traditional healthcare use amongst those dying of HIV/AIDS and other causes in the Agincourt HDSS study site from paper IV.
Table 3: Analysis of Age and Gender standardised AIDS/TB mortality rates and rate ratios aggregated by village for 2007-2008 and 2009-2010 (Supplemental materials for Paper 2)

<table>
<thead>
<tr>
<th>Village</th>
<th>2007-2008</th>
<th>2009-2010</th>
<th>2009-2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths</td>
<td>Person Years</td>
<td>Standardised Mortality Rate / 1000 PY</td>
</tr>
<tr>
<td>A</td>
<td>74</td>
<td>10169</td>
<td>7.34</td>
</tr>
<tr>
<td>B</td>
<td>34</td>
<td>4634</td>
<td>7.41</td>
</tr>
<tr>
<td>C</td>
<td>45</td>
<td>7717</td>
<td>6.12</td>
</tr>
<tr>
<td>D</td>
<td>30</td>
<td>5126</td>
<td>5.90</td>
</tr>
<tr>
<td>E</td>
<td>22</td>
<td>3886</td>
<td>5.93</td>
</tr>
<tr>
<td>F</td>
<td>40</td>
<td>5255</td>
<td>7.53</td>
</tr>
<tr>
<td>G</td>
<td>21</td>
<td>3423</td>
<td>6.22</td>
</tr>
<tr>
<td>H</td>
<td>62</td>
<td>7022</td>
<td>9.21</td>
</tr>
<tr>
<td>I</td>
<td>42</td>
<td>6704</td>
<td>6.37</td>
</tr>
<tr>
<td>J</td>
<td>40</td>
<td>6333</td>
<td>6.39</td>
</tr>
<tr>
<td>K</td>
<td>37</td>
<td>9870</td>
<td>7.90</td>
</tr>
<tr>
<td>L</td>
<td>23</td>
<td>3058</td>
<td>7.87</td>
</tr>
<tr>
<td>M</td>
<td>41</td>
<td>4707</td>
<td>8.94</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>2596</td>
<td>7.97</td>
</tr>
<tr>
<td>O</td>
<td>39</td>
<td>4165</td>
<td>9.59</td>
</tr>
<tr>
<td>P</td>
<td>37</td>
<td>5882</td>
<td>6.34</td>
</tr>
<tr>
<td>Q</td>
<td>16</td>
<td>2633</td>
<td>5.97</td>
</tr>
<tr>
<td>R</td>
<td>10</td>
<td>1490</td>
<td>7.15</td>
</tr>
<tr>
<td>S</td>
<td>11</td>
<td>1414</td>
<td>7.93</td>
</tr>
<tr>
<td>T</td>
<td>9</td>
<td>1229</td>
<td>6.90</td>
</tr>
<tr>
<td>U</td>
<td>14</td>
<td>2123</td>
<td>6.72</td>
</tr>
<tr>
<td>V</td>
<td>29</td>
<td>4459</td>
<td>7.01</td>
</tr>
<tr>
<td>W</td>
<td>12</td>
<td>1724</td>
<td>8.03</td>
</tr>
<tr>
<td>X</td>
<td>14</td>
<td>1627</td>
<td>8.63</td>
</tr>
<tr>
<td>Y</td>
<td>6</td>
<td>823</td>
<td>0.39</td>
</tr>
<tr>
<td>Overall</td>
<td>728</td>
<td>107980</td>
<td>6.88</td>
</tr>
</tbody>
</table>

PY = Person years of observation
DISCUSSION

The series of studies which constitute this dissertation seek to understand the details of the development of a localised HIV epidemic in a rural South African community and the manner in which individuals adapted their health-seeking behaviour in response to the epidemic. There is a particular focus on the period of time over which ART was introduced. Whilst this is of historical interest it also has the potential to inform those responsible for developing HIV testing and treatment provision about the ways in which these programmes may be optimised to maximise their effectiveness.

Paper I confirms previous explanations about how the epidemic developed in South Africa. Deaths due to HIV/AIDS were seen from the earliest days of the epidemic to be distributed in all villages throughout the study site. This fits with the idea that HIV was introduced to rural areas of South Africa by migrant workers, mainly employed in the mining industry, returning home to their families in the rural areas and infecting their partners who lived in those communities. Importantly, there is no evidence that it was introduced initially by particular sub-groups of the population. Studies in rural KwaZulu-Natal by Coffee et al. [7] and Lurie et al. [42] indicate that although migration was a major driver of the introduction of HIV to these rural communities, the subsequent spread was through both males and females in the community forming multiple sexual relationships. A similar pattern would explain the subsequent rapid spread of the epidemic in the Agincourt community. There was some evidence that at least one community failed to see a reduction in HIV/AIDS mortality subsequent to the introduction of ART in the area.

A comparison of HIV prevalence data and HIV/AIDS mortality patterns in this paper shows that at the peak of the epidemic women aged less than 40 had higher rates of HIV/AIDS mortality than males of the same age group, which correlates well with the differences in HIV prevalence between the two genders as recorded nationally in 2011 [158]. Age disparate sexual mixing patterns in which younger females were in relationship with older males have been proposed as a major factor explaining this. In the Agincourt HDSS study site the differences in HIV-related mortality rates between males and females aged less than 40 have decreased over time. A possible explanation for this is that these younger women have been able to take greater advantage of the availability of ART in this community since 2007 than their male contemporaries so improving the chances of long term survival post-infection.

In the older age groups HIV prevalence was similar for both genders yet the levels of male HIV/AIDS mortality were much higher than that for females. This may be explained by the evidence from the South African 2012 HIV prevalence survey of the lower likelihood of men getting tested for HIV compared with women and the lower probability that they will be receiving treatment when HIV positive [22].

The true economic and social impact of the epidemic becomes clear from the estimate that over 60% of the deaths occurring from 1992 to 2013 were attributable to HIV. As those at most risk of infection are younger adults this has led to a significant shift in age-related mortality
patterns. Consequences of this include increased levels of poverty amongst the affected households [44] and the emergence of households either headed by orphaned children or in which children are raised by grandparents [159, 160].

Paper II investigated spatial patterns of HIV/AIDS mortality amongst adults over the period that ART was rolled out in this community. In this study, rates were aggregated by either village or 2km-squared grids in order to assess whether there was evidence for different epidemic trajectories in different geographical regions of the site. For one village the HIV-related mortality rate increased over the period that a 30 % decrease occurred across the whole site. As the rates were standardised by age and gender, other factors must explain these figures. Localised differences are also seen in the gridded analysis, particularly towards the extreme east of the site. The increase in the rate for the grid squares overlapping the Bhubezi Community Health Centre suggests that the distance individuals lived from the health centre appear not to be a main determinant of the risk of dying of HIV.

These results provide further evidence for the importance of understanding the characteristics of local HIV epidemics in order to ensure that the limited resources available for prevention, testing and treatment activities are used most effectively. A UNAIDS report on the local HIV epidemics emphasises the potential use of detailed small-area HIV prevalence data to locate key population sub-groups with a particularly high HIV prevalence [152]. This definition often includes groups such as commercial sex workers, men who have sex with men, and injecting drug users. In the context of an area such as the Agincourt HDSS study site this could be extended to include groups such as younger women who have a significantly heightened risk of HIV acquisition compared to the general population. It is increasingly recognised that if we are to reduce HIV rates to low endemic levels new approaches are needed to support effective prevention, testing and treatment for these groups. Results of studies in rural [146] and urban [148, 149, 151] areas of KwaZulu-Natal, South Africa, which show evidence for local HIV hotspots within a more generalised epidemic, further emphasise this.

An important difference between this and other spatial epidemiologic analyses of the HIV epidemic, to which the results of this analysis are compared, is that HIV/AIDS mortality rather than HIV prevalence is used as the outcome measure. For an individual to die of an HIV related illness they must first become infected and the illness must subsequently progress to death. An individual’s risk of death would be contributed to by the risks of unsuccessful outcomes at each stage (e.g. late testing, late ART initiation and failure of or drop-out from treatment). Having only the final outcome makes it difficult to make causal inferences at each preceding stage.

The results from Paper III help us to understand more details of the factors that may explain the spatial differences in the mortality patterns that were seen in Paper II. It shows that an individual’s risk of dying of HIV/AIDS over the period of the ART roll-out was determined by the interplay of a combination of risk factors. The study found that whilst the decrease in mortality was much greater for those dying of HIV/AIDS than those dying of other causes, many of the factors associated with this risk acted in the same manner in both groups. This
suggests that the model in paper III fails to capture some of the important determinants which explain the decrease in HIV/AIDS mortality. However this is also potentially an important finding as it suggests that interventions which address these determinants, such as structural measures designed to alleviate poverty or health systems interventions which are designed to lower barriers to access, will benefit all of those in the community, not just those who are HIV infected. It would be interesting to explore in future studies whether a variable describing the changes in availability of ART over time could be included and whether this would improve the fit of the model.

As was discussed previously, structural factors in South African society are thought by many to have been a key driver for the development of the HIV epidemic. The results from Paper II would seem to confirm this. Relative poverty, as measured by asset based SEP, is seen to be a key determinant of risk. The level of educational attainment is also an important risk factor. However as the two variables remain independently predictive when combined in a multivariate model, this suggests that educational attainment rather than being a proxy for household wealth, captures other dimensions of risk. These may include the fact that at least for young women, attending school regularly during the years at which their risk of infection is particularly high may serve protect them from engagement in high risk sexual encounters with older men. Also successful completion of education may open up employment opportunities which increase their empowerment and may lead to a reduction in HIV risk behaviour. This has been discussed in detail by other authors [161].

As discussed previously many studies of determinants of HIV risk use HIV prevalence, or the risk of being infected, rather than HIV/AIDS mortality as an outcome. It is therefore useful to compare the associations seen here with those studies. A 2011 study of a representative sample of the Agincourt HDSS study site population reported on the associations between HIV prevalence and socio-demographic characteristics [158]. It was found that the HIV prevalence was lower amongst those of higher SEP and higher levels of education. This is consistent with the findings of this study. The findings on the protective effect of higher levels of education are consistent with those of Bärnighausen et al. in rural KwaZulu-Natal [55] which found that each additional year of education for an individual conferred increased protection against the risk of infection. In order to better understand the associations between HIV mortality and SEP it would be useful to develop a more fine-grained model of causality in which associations at each stage between initial infection and death are understood.

Other studies have shown that the association may be context dependent. Hargreaves et al. suggested that the relationship between HIV and SEP appears to be changing over time [49]. They present evidence that in the earlier stages of the epidemic those of higher SEP were at an increased risk of infection, whilst later higher SEP appears to be protective. Following this hypothesis would lead us to the assumption that the HIV epidemic in the Agincourt HDSS study site is relatively mature [53]. As this study only analysed data over four years it was unlikely to identify significant changes in SEP so it is not possible to draw conclusions about temporal trends.
Competing factors appear to explain the associations identified between gender and the risk of HIV/AIDS mortality. Younger women are exposed to a significantly greater risk of infection than their male counterparts, due to their engagement in sexual relationships with older men [74], their higher risk of infection per sexual act [162] and their exposure to gender violence [163]. However this is balanced to some extent by the protective effect of better health seeking behaviour. Data from the 2012 South African national HIV prevalence, incidence and behaviour survey show that women are more likely to be tested and to be on treatment when HIV-positive compared to their male counterparts [22]. In the Agincourt HDSS study site it appears that at the peak of the epidemic the increased risk of infection for younger women outweighed any protective effects of better health seeking behaviour. However, by 2013 the mortality rates for younger women appear to converge with those for younger men (Paper I) indicating that the balance between these competing factors had changed.

There was some evidence of an association between the distance individuals resided by road from the Bhubezi Community Health Centre and the risk of HIV/AIDS mortality. Those living greater than five km away were at a greater risk of death due to HIV/AIDS when controlling for other parameters in the model. This pattern is consistent with the finding of Cooke et al. [57] of a strong negative association between the distance individuals lived from a clinic providing treatment and the uptake of ART. However the spatial analysis (Paper II) also indicated that there were hotspots of high mortality close to the clinic.

An interesting negative result from paper III was the lack of an association between the taxi fare to reach the Bhubezi Community Health Centre and HIV/AIDS mortality. It was hypothesised that for the cost of travel by public transport would present a barrier to regular clinic attendance thus providing an explanatory mechanism of the relationship between wealth and an individual’s risk of dying of an HIV/AIDS. This appears not to be the case, possibly due to the availability of social grants as a means to overcome absolute poverty.

In the multivariate model developed in paper III a variable representing whether or not a household owned motorised transport showed an independent protective effect when controlling for the effect of household socioeconomic position (SEP). A possible explanation for this is that the motorised transport is used to travel to clinic visits. An alternative explanation is that this variable is a proxy for the cash wealth rather than asset wealth available to a household. This cash wealth may enable the household to overcome the financial shock of the loss of income when a member is seriously ill, thus ensuring they have access to better nutrition and healthcare.

The finding in the multivariate model that having a Mozambican country of origin had a protective effect differs from results of earlier studies in this community. Hargreaves et al. reporting on data from 1992 to 2010 [164] found that mortality for children aged between one and five years was higher for those from formerly Mozambican households than for those from South African ones. The results in Paper I indicate that in the earlier stages of the epidemic HIV-related mortality rates were higher for those of Mozambican origin than South Africans but in more recent years these rates have converged. Polzer-Ngwato [134] identified an
ongoing process of assimilation through which the cultural differences between those who families arrived as refugees from Mozambique and those in the host South African community decreased over time. Williams et al. also identified a convergence of fertility patterns between females of Mozambican origin and South Africans between 2003 and 2009 [165]. In 2011 it was found that in this area the HIV prevalence of those of Mozambican origin was actually lower than that for South Africans. This evidence for a protective effect in the most recent data suggests that certain characteristics of the former Mozambican community may lead to a decrease in the risk of HIV/AIDS mortality. This may be linked to behavioural factors, such as sexual mixing patterns, however more detailed qualitative work would be needed to confirm this.

Papers II and III addressed the quantitative question of ‘what’ happened in relation to general and cause specific mortality over the period of the roll-out of ART in this community. Paper IV is an attempt to understand more about changing patterns of healthcare seeking behaviour over this period which may have contributed to these changes. We see that between 2003 and 2011 there was a decrease in the use of traditional forms of healthcare and an increase in the use of biomedical healthcare amongst those dying of any cause. For those who subsequently died of HIV/AIDS, the use of traditional healthcare was higher in the earliest of the three periods studied (2003-2005) which is likely to reflect the lack of easily accessible alternative treatment options at that time. Over the subsequent periods, 2006-2008 and 2009-2011, as ART became more easily accessible we see the same pattern of change as seen for all causes of death. However the dual use of traditional and biomedical care during the final illness remains high.

We can conclude from this that even though the previous barriers to accessing ART have been lowered, traditional healthcare is still widely used by those living with HIV in this community. There are potential benefits that could come from a greater integration of traditional and biomedical providers of care. The traditional providers are likely to have a high level of cultural sensitivity and provide effective care for psycho-social aspects of HIV related illness, which will typically be beyond the scope of biomedical providers [166-168]. However there is some evidence that patients taking herbal treatments in conjunction with ART may delay treatment initiation, and have lower levels of adherence or take gaps in treatment [169, 170]. There is also evidence for possible antagonistic interactions between components of ART and certain traditional remedies [171]. Anecdotal evidence from Agincourt HDSS fieldworkers indicates that as those on ART are told categorically to stop taking traditional treatments, much of this use is hidden from biomedical practitioners for fear of disapproval and potentially prejudicial treatment by the healthcare providers. There is however evidence from recent studies in Mozambique that the referral of patients from traditional healers to biomedical facilities may be improved if adequate training is given [51, 172].

This longitudinal study provides quantitative evidence for the finding from earlier qualitative studies carried out in this community that a somewhat sophisticated evidence-based healthcare marketplace exists [85, 86]. Assuming this to be the case we can conclude that the evidence for
the efficacy of ART has led to dramatic changes in patterns of healthcare use. Further qualitative research is needed to understand the changing patterns of healthcare use in communities such as this in more detail. A major challenge over the coming decades will be maintaining increasing numbers of patients on ART with adequate levels of adherence in order to prevent treatment failure and the development of resistance to the drugs used. In depth qualitative studies which aim to understand the factors determining patterns of health-seeking behaviour will be important in understanding how to achieve this.

The Agincourt HDSS has been carrying out a continuous population survey collecting socio-demographic and mortality data in this rural community for over 20 years. As a result the dataset provides what may be a unique insight into the impact of a local HIV epidemic from its earliest days when the first cases are likely to have occurred. The use of this dataset is a major strength of the study allowing us to better understand the entire course of the epidemic to date and also focus in on the time period when dramatic reversals in mortality trends were seen, coincidental with local provision of ART.

Using mortality data we cannot probe specific questions such as, ‘What are the risk factors associated with acquiring HIV in this community and how have those changed over the course of the epidemic?’ or ‘What are the relationships between changing patterns of HIV testing and changes in stigma associated with HIV?’. We would also ideally like to understand in more detail the complex process of engagement with healthcare providers and understand more of the differences between those who are HIV infected and successfully treated and those who default or for whom treatment in unsuccessful. Gaining a better understanding of these processes through both quantitative longitudinal studies and in-depth qualitative research will be essential to reaching the goal of reducing the HIV epidemic to levels of low endemicity and eventual elimination [163, 173, 174].

In these studies a person’s place of residence is used as the geographic determinant of mortality risk. In a community such as this people are exposed to risk in many different places. A large proportion of the working age population are economic migrants moving between their rural home and employment often in urban areas. Consequently they are hence exposed to a very different HIV risk environment than that found in the rural community for much of the time. To some extent this is captured in our models through the inclusion of variables representing the proportion of time they are resident in the different locations, but we would ideally wish to represent the changing spatial dynamics associated with migration in a more detailed manner in our models.

The relationship between proximity to healthcare and the eventual outcomes is complex. In a highly stigmatised condition such as HIV it may be that the ease of accessing one’s local clinic is not the most relevant factor but rather the barriers to attending one’s healthcare provider of choice. In the analysis for paper III the assumption is made that during the period 2007-2010 those who needed treatment for HIV related illnesses were likely to use the Bhubezi Community Health Centre rather than other clinics in the area. However it is clear from paper IV and other studies [175] that complex pluralistic sources of healthcare were available and
There is evidence from the VA narratives that many patients arrived at Bhubezi after exhausting other traditional and biomedical treatment options or because they were seeking a second opinion after already having received a diagnosis in a public clinic [unpublished data].

The analysis in paper IV of patterns of healthcare use amongst the subset of the population infected with HIV who subsequently died of HIV-related illnesses has to be carefully interpreted. In the period before ART was easily accessible there was an inevitable progression between infection and eventual death. Hence patterns of healthcare use amongst those who die are likely to be largely representative of patterns amongst all those infected. However in the period from 2007 onwards, when ART was more easily accessible, we are excluding from the study population those who are successfully treated. Thus the reported effect size for the switch from traditional to biomedical care reported over this period is likely to be an underestimate of the true level we would measure in the general population, if such data were available.

It is important to gain a more nuanced understanding of the types of traditional healthcare used. Traditional treatments may be targeted as being directly curative or supportive, alleviating symptoms and side effects of disease. They may also focus on spiritual and psychosocial elements of illness. Further research is needed to better understand the way in which individuals interact with different forms of healthcare. This may facilitate more effective cooperation between traditional and biomedical providers.

In any longitudinal survey, particularly one using proxy informants to report on key events that have happened between visits, there is the potential for recall bias. Detailed quality control processes have been developed in the Agincourt HDSS to ensure that data collected are consistent with past reports and in addition ongoing processes of retrospective data cleaning are carried out. VA interviews are carried out as soon as possible after a death is notified, but as census visits occur only once a year it could be that family members are asked to recall signs and symptoms of illness that occurred many months previously. The use of a specifically trained team of verbal autopsy fieldworkers is designed to minimise the effect of this and also manage the emotional impact on respondents of recalling often traumatic past events. A study from 2009 by Byass et al. using data collected in Burkina Faso and Indonesia showed that when VA’s were repeated a significant variation in the interview responses occurred [176]. This in turn led to individual level discrepancies in the causes of death assigned.

The potential bias introduced due to the occurrence of deaths for which no cause could be assigned, often due to a lack of reported symptomatic data, is addressed in particular in papers II and III. It appears that there are more missing data in certain age/gender strata such as younger working men due perhaps to the time spent away from the area during the period leading up to their death. As it is known that these sub-groups are also at a heightened risk of HIV infection the results may to some extent underestimate the true levels of HIV/AIDS mortality.

It has been shown that the true impact of HIV on population level mortality is much greater than that which would be seen by assigning all deaths as HIV-related or not based on the most
likely cause. Slaymaker et al. have developed estimates of HIV attributable mortality based on data from longitudinal HIV cohorts in a number of Sub-Saharan African countries [177]. These estimates when applied to the Agincourt HDSS study site population data making reasonable assumptions about historical HIV prevalence led to an estimation that over 60 % of the adult deaths occurring between 1992 and 2013 were attributable to HIV infection and hence would not have occurred at the time that they did in its absence (Paper I). It is also possible to calculate the relative increase in likelihood of an HIV infected person dying of a particular cause compared with their HIV negative counterparts [127]. For example the mortality rate ratio comparing those HIV positive and HIV negative for death from acute respiratory infection is close to 40. It may be argued based on this that the more cases should be included in the ‘HIV/AIDS’ category. This is likely to affect the associations and mortality patterns seen in this study.

Another perhaps more fundamental limitation of analyses such as these is that all individuals are treated as isolated units. To some extend this has been addressed in the regression models by clustering individuals by household (based on the assumption that they have more in common with household co-residents than others). However this still fails to take account of the fact that individuals exist in complex social networks including family and non-family members [178, 179]. The nature of these networks and the extent to which they are either protective or expose individuals to greater risk are not represented and are likely to have an important influence on risk associated behaviour.

An important question which arises from a series of studies such as this and indeed any study from a relatively small sampled population such as that in an HDSS site is whether these results may be generalised to other contexts [97]. To some extent in South Africa the legacy of Apartheid has created a unique social structure which differs from that in other countries of the region and the wider area beyond. However parallels can be identified, certainly in countries emerging from the legacy of colonialism and experiencing the effects of rapid change due to industrialisation and globalisation [38]. The context thought to have driven the epidemic, the social dislocation resulting from forced labour migration, is common to many rural areas within the country and in the surrounding region. There are also likely to be particular cultural factors which make the Agincourt HDSS study site, in which around a third of the population are either former Mozambican refugees or their families, different to other rural communities. Paper III does provide evidence that despite the assimilation into the host community over time there are still important differences between the groups that may affect health outcomes [134].
CONCLUSIONS

In 2015 we are at a critical moment in the global battle to combat the HIV/AIDS epidemic. Since 2004 there have been significant decreases in HIV/AIDS-related mortality throughout the world, particularly so in the worst affected regions such as Sub-Saharan Africa. For example in South Africa, the median age of death fell to 42.6 years in 2004 but by 2013 returned to 52.5 years, the level seen prior to the HIV epidemic [163]. These dramatic changes are largely due to the success of programmes to provide life-saving ART drugs. In 2003 less than a million people worldwide were receiving ART, the vast majority of these in high income countries, by 2013 this had risen to 13 million, close to 10 million of whom were in the African region [163]. This has been achieved through remarkable efforts of governments in the worst affected regions, supported by global funding bodies, to roll out the treatment and ensure it is available at an affordable price. Despite these successes the situation is still grave. In 2014 there were close to 5,500 people newly infected with HIV each day worldwide, and close to 1,000 of these were in South Africa [1]. Whilst some significant breakthroughs are being made in understanding the biology of HIV, which it is hoped will eventually lead to more effective treatments and ultimately a vaccine, we are still many years away from seeing these become available. Modelling studies indicate that, using currently available approaches, the HIV epidemic can be reduced to endemic levels by 2030 [163, 174, 180]. This will require a scale-up in investment and effort from current levels to achieve these ambitious targets. The major concern amongst those leading these efforts is around complacency in the international community, due to a view that the AIDS problem has been ‘dealt with’. As can be seen from these figures this is far from the case; indeed a failure to scale up the response could lead to a reversal of the achievements made to date.

A key question in summing up this work is how studies such as this may contribute to these efforts. In 2007 a report released by UNAIDS emphasised the crucial importance of countries developing the capacity to monitor and understand the nature of their epidemic and the impact of the response that has been mounted [181, 182]. Subsequent reports have emphasised the importance of having detailed information about local epidemics in addition to understanding the situation at a national or international level [152]. Such studies have been shown to reveal significant local variations in the trajectory of the epidemic. These may be due to differential access to testing and treatment, due to physical or socio-economic barriers to accessing healthcare or other causes. They may also be due to the presence of sub-groups within the population who may be hard to reach by those providing HIV testing and treatment services. Currently much of the measurement of epidemic trends relies on modeling studies using the available data and reasonable assumptions to extrapolate that data to a wider area. The availability of more detailed and representative primary data will lead to an improvement in the accuracy of these models.

Studies such as this, analysing mortality trends in a local area, can be used to develop analytical methods that health planners may use to better understand national level health trends. The detailed information gained on the development of local epidemics through the
continuous monitoring of a geographically defined cohort as described in this study is useful for validating conclusions that come from studies across wider areas, such as DHS surveys.

Monitoring epidemic trends using mortality outcomes alone is a rather blunt instrument. Moving forward, more detailed longitudinal studies are needed to better understand how the factors associated with increased mortality risk operate at each stage of the HIV testing and treatment continuum [173, 183]. It may be, for instance, that those in this community who are employed and hence relatively wealthy present for testing and treatment at a later stage in the progression of their illness due to the opportunity costs associated with the time needed to attend clinic appointments. At the same time their long term outcomes may be better than those who are poorer, perhaps due better levels of nutrition and hygiene.

The quantitative studies included in this dissertation address the question of ‘What has happened?’ in relation to the HIV epidemic. There is an urgent need for more qualitative studies which address the question of ‘Why have these things happened?’ Gaining such understanding of the underlying aspects of human behaviour will be essential for the development of more effective testing and treatment programmes. For example, the mortality trends suggest that one community in the study area has failed to benefit from the provision of ART whilst other neighbouring villages have seen major improvements. Through in depth interviews with community members we may gain insights into the underlying causes of these changes which could be used to address the particular barriers to healthcare access. The same approach is needed to better understand how relative poverty or a lack of education leads to poorer health outcomes in a situation where HIV testing and treatment is provided free at the point of care. Gaining this insight will enable a better understanding of how to address the structural factors in this society which appear to be the major drivers of differential mortality outcomes.
ACKNOWLEDGEMENTS

It’s easy to reduce the HIV/AIDS epidemic to numbers in tables rather than the faces and names. I have, perhaps inevitably, been somewhat insulated from the awful human impact of HIV in the community I have written about. However I have had the privilege of knowing friends who were brought back from serious illness and restored to vitality after getting on treatment and sadly others whose lives were cut tragically short. For obvious reasons they will remain anonymous but their stories have been a motivation to make this piece of work as useful as it can be.

This PhD has led me on a remarkable journey intellectually, professionally and geographically. So many people have contributed to that journey and inevitably in identifying particular individuals I wish to thank, I will miss out others.

Without the data on which all of this work is based none of this would have been possible so my deepest gratitude goes to all of the fieldworkers and data staff working with the Agincourt HDSS for their professionalism and commitment. You are a remarkable team and it has been an honour to work with you.

I would like to thank Kathleen, my wife and constant supporter throughout this enterprise, for carrying the burden of domestic and family life so often whilst I was closeted away and also for the insights into the human face of the epidemic she brought back from the ‘coal-face’ serving this community in a rural South African hospital. I really couldn’t have done this without you. Part way through our time in South Africa our children, Daniel Jude and Vivienne Grace were born, and their sense of love, joy and mischief have helped me keep a sense of perspective throughout.

My parents John and Barbara my sister Alison and my mother-in-law Rosemarie have supported us all in so many ways throughout, despite the separation when we were in South Africa. I hope the product of this work helps to explain something of the importance of it all.

It’s been a privilege to have as my supervisor from Umeå University Professor Peter Byass. He opened the door for my PhD studies and kept it held open providing friendship, support, encouragement and insight throughout the process. The hospitality from Peter and his wife Margaret in the UK, Sweden and South Africa to myself and my family has been a wonderful bonus.

Over the last decade I have had the privilege to work twice with the MRC/Wits Agincourt research unit. This is a remarkable organisation and group of individuals who have dedicated themselves to serving this community over many years. Steve Tollman and Kathy Kahn have been the visionary leaders of this project since its inception. I have learnt so much from both of them about what is involved in sustaining a project such as this whilst striving for and achieving scientific excellence. They have achieved all this through the development of long term working relationships and friendships with world-leading scientists from which I have benefitted immensely. Amongst this group I would particularly like to thank Jane Menken and
her team from the University of Colorado (Boulder), Margaret Thorogood and Gillian Hundt from the University of Warwick, Charles Newton from the University of Oxford and Michael White from Brown University, who have all provided much practical support and guidance along the way.

Mark Collinson has been a friend, mentor and latterly PhD co-supervisor. I have greatly benefitted from his insight into this community gained over many years spent living and working in rural Bushbuckridge. His encouragement and insightful critique have immensely improved this work. Mark’s wife Alison also provided much friendship and hospitality during our time in South Africa.

Sangeetha Madhavan, my other co-supervisor, has worked with and at times lived amongst this research community over many years. Her cultural insight and ability to link the big picture with a focus on the individual stories that make it up has been a great help. She provided insight and wise counsel through the development of the work.

Within the MRC/Wits Agincourt unit I’ve had the pleasure of working with and living alongside many colleagues and friends. By its nature working for the unit blurs the boundaries between working and social relationships in a wonderfully creative way, with your friends and neighbours being your colleagues and co-investigators.

Xavier Gómez-Olivé has amongst other things he taught me many important lessons about leadership. Rhian Twine, through her work in ensuring the scientific work is relevant, acceptable to and to the benefit of our research community, has consistently challenged me in a very positive way to follow her lead. It’s also been a real pleasure to work and study alongside Ryan Wagner, my fellow Umeå PhD student.

My colleagues in the data section, Chodziwadziwa Kabudula, Afolabi Sulaimon, Pleasure Ngobeni, Carlos Horylka and Itayi Adam taught me so much technically and helped me understand African (and Brazilian) culture. I’d also like to mention Kobus Herbst, Sam Clark and Benjamin Clark who built many of the technical foundations for the Agincourt data systems which we and so many others have benefitted from.

I also have had much support way above and beyond the call of duty from the team based at the School of Public Health at the University of the Witwatersrand in Johannesburg in particular Jackie Roseleur, Dawn Dalby, Wendy Pearsall, Lynda Oldert, Sadiya Ooni, Melta Buthelezi and Nicolette Pingo. Thanks to you all.

Amongst the Agincourt field staff and all members of our research team I would particularly like to thank Jeffrey Tibane, Simon Khosa, Bernard Silane, Sizzy Ngobeni and Busi Ngcobo for giving of their time and wisdom so freely to give me insights into the local Shangaan culture to help me make some sense of the findings from this work. The MRC/Wits Agincourt unit could not exist without the dedicated administrators and support team who quietly and efficiently carry a huge administrative burden in particular; Violet Chela, Doreen Nkuna, Ngoni Ngwarai and Floidy Wafawanaka.
To all of my co-authors, your formative ideas, reviews and critiques have ensured that all these papers are immensely better than they would otherwise have been.

To all the staff and residents at Wits Rural Facility, our South African home, thanks for creating such an amazing place to work and think. In particular I’d like to thank David Bunn & Melissa McHale and Sharon Pollard & Kevin Mitchell our neighbours and friends for their inspiration, support and encouragement.

To my friends Ronald Mathebula and Charity Malinda for making us know that Bushbuckridge was a community that myself and my family really belonged to, your love support and inspiration have meant a lot to us. Also to Steve Mosehla and the other members of the mighty Klaserie Hawks Athletics Club, I am looking forward to pounding more South African roads with you all in the years to come.

Tara Polzer, Paul Pronyk, Julia Kim, Mosa Moshabela and many others in different ways laid the foundation for this work from their own lives, work and research in the Bushbuckridge area. I am grateful for your support both directly and indirectly.

Among my fellow researchers in the School of Public Health at the University of the Witwatersrand I would like to particularly thank Benn Sartorius, the ultimate shape of this PhD owes much to his early guidance on the structure of the core papers. Also much of this analysis builds upon his and Eustasius Musenge’s prior work.

Thanks also to Colby and Ann Gottert for generously allowing me to use some of their photographs. Theirs are the lower picture on p11, all those on p19, the lower picture on p27, the lower picture on p28 and the middle picture on p45.

The completion of this PhD has been a period of transition returning to the UK from South Africa. In my new academic home at the London School of Hygiene and Tropical Medicine I have been grateful to my new friends and colleagues in the PRIME team. In particular; James Hargreaves, Calum Davey, Bernadette Hensen, Shari Krishnaratne, Stephanie Dringus, Jennifer Thompson and Liz Fearon for welcoming me into your group and keeping me focussed on where I was heading.

It’s been a real pleasure to build new relationships with colleagues in the Umeå Centre for Global Health Research as a result of my visits to study and work there. In particular Stig Wall the founder of the unit and long term friend and supporter of the Agincourt unit, Anneli Ivarson, John Kinsman, Nawi Ng, Miguel San Sebastian, Lucia d’Ambruoso, Karin Johansson, Lena Mustonen and Birgitta Aström. I would also like to thank in particular Kerstin Edin who has provided wonderful hospitality and fantastic meals during my trips to Umeå and her trips to South Africa.

Finally I would like to again pay tribute to the research community in the Agincourt sub-district for their ongoing support and willing sacrifice of time over many years that has allowed this work to be possible.
None of this would have been achieved without the funding for the project provided by a range of organisations over many years. The University of the Witwatersrand, Medical Research Council and Mpumalanga Province, South Africa have provided support since the inception of the Agincourt HDSS. The Wellcome Trust, UK has provided core funding for well over a decade (Grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z). I also acknowledge important contributions from The William and Flora Hewlett Foundation, National Institute on Aging (NIA) of the NIH, The Andrew W Mellon Foundation, USA and the Anglo-American chairmans fund. This work was carried out in collaboration with the Umeå Centre for Global Health Research, supported by Forte, the Swedish Research Council for Health, Working Life and Welfare (grant no. 2006–1512), and the European Union International Research Staff Exchange Scheme (grant no. 295168).
REFERENCES


61. Samsom M MK, va Niekerk I: Social Grants in South Africa : ODI Policy brief
2005.

62. Leclerc-Madlala S: 'We will eat when I get the grant': negotiating AIDS, poverty and

63. Hardy C, Richter M: Disability grants or antiretrovirals? A quandary for people with

64. WHO: Scaling Up Anti-Retroviral Therapy in Resource Limited Settings A Public Health

65. WHO: Treating 3 million by 2005: making it happen: the WHO strategy: the WHO and
UNAIDS global initiative to provide antiretroviral therapy to 3 million people with

treatment outcomes and household well-being in KwaZulu-Natal. Development Southern


68. UNAIDS report on the global AIDS epidemic 2012

69. UNAIDS: Agenda for Accelerated Country Action for Women, Girls, Gender Equality
and HIV
http://www.unaids.org/sites/default/files/media_asset/20100226_jc1794_agenda_for_accele
rated_country_action_en_0.pdf. 2010.

70. Watts C, Seeley J: Addressing gender inequality and intimate partner violence as critical
barriers to an effective HIV response in sub-Saharan Africa. Journal of the International
AIDS Society 2014, 17(1).


72. MacPhail C, Williams BG, Campbell C: Relative risk of HIV infection among young men
and women in a South African township. International journal of STD & AIDS 2002,
13(5):331-342.

73. Maughan-Brown B, Kenyon C, Lurie MN: Partner Age Differences and Concurrency in
South Africa: Implications for HIV-Infection Risk Among Young Women. AIDS and
Behavior 2014, 18(12):2469-2476.

74. Ott MQ, Bärnighausen T, Tanser F, Lurie MN, Newell M-L: Age-gaps in sexual


99. The demographic and health surveys program [http://www.measuredhs.com/]


104. [http://www.indepth-network.org/]


