# CHAPTER 1.0: INTRODUCTION

## 1.1 Background

More than 65 million people have been infected with human immunodeficiency virus (HIV) worldwide and more than 25 million have died of acquired immunodeficiency syndrome (AIDS) associated conditions since the HIV virus was first reported more than two decades ago [1, 2]. In the African continent it was reported that HIV/AIDS is the leading cause of mortality and morbidity [3].


**Figure 1**Global estimation of persons living with HIV/AIDS in 2005 (millions); total 40.3 (36.7-45.3) million. (Source: UNAID/WHO, 2005)

In South Africa, the HIV/AIDS epidemic has had a major socioeconomic impact particularly in high prevalence areas. South Africa has more people living with HIV/AIDS than any other country in the world and faces enormous challenges in scaling up its response to the now-mature and generalized HIV/AIDS epidemic [4]. The magnitude and growth of the prevalence rates of HIV infection differ by province. The highest HIV prevalence rates are in KwaZulu Natal (KZN) province at 37.5% among pregnant women, followed by Mpumalanga and Free State [4]. This high prevalence has resulted in high AIDS mortality among the residents. There is evidence in the literature that HIV/AIDS epidemic has also had a substantial impact on the country's overall social and economic progress. For instance, average life expectancy in South Africa is now 54 years – without AIDS, it is estimated that it would be 64. Over half of 15 year olds are not expected to reach the age of 60 [5].

The Africa Center Health and Demographic Surveillance Site (Africa Centre HDSS), in rural KZN province of South Africa is one of the well established HDSS in the African continent. All data on mortality (causes of deaths) were collected via the verbal autopsy (VA) approach, and these data were collected by trained and experienced filed officers. HIV/AIDS with or without Tuberculosis (TB) remains the major cause of death among the Africa Centre HDSS population [55] and residents are at a high risk of HIV/AIDS. HIV/AIDS knows no boundaries, but epidemiology using a sophisticated geographic information systems (GIS) analysis may yield vital clues as to HIV distribution (clustered or random) and mortality across communities with different norms and socio-economic status. In light of the HIV/AIDS burden especially in the rural areas of South Africa it is now becoming necessary to look at new management approaches to help manage HIV/AIDS at the community level (district level) and GIS applications can be one of the appropriate approaches in terms of targeting high risk areas for intervention.

In this study our aim is to employ the applications of GIS and a Kulldorff's spatial scan statistic (spatial clustering) to investigate statistically significant mortality clusters ($p \leq 0.05$) of causes of deaths to detect areas of significantly high or low mortality (rates) in the Africa Centre HDSS population at the community level, so appropriate health intervention can follow to help manage HIV/AIDS. Spatial clustering of AIDS-related mortality has not been done before at the community level so this study was the first to address this using the Africa Centre demographic information systems (ACDIS) database. To effectively address the menace of HIV/AIDS mortality, it is important to know the spatial distribution of deaths so that health intervention programmes can strategically target these clusters. Statistical methodology to identify disease cluster is consistently under review and development. Kulldorff's spatial clustering method was applied to identify statistically significant AIDS mortality clusters and to compare spatial clustering patterns of AIDS mortality in adults with adult all-cause mortality, other cause of death categories and all-cause childhood mortality (<5 years).

**1.2 Objectives**

The objectives of this study were:

1. To use GIS and spatial scan cluster detection method to identify statistically significant AIDS mortality clusters ($p \leq 0.05$) between 2000-2006
2. To compare the spatial clustering patterns of AIDS mortality in adults with:
    a. Adult all-cause mortality

    b. Other cause of death categories (injury, non-communicable, communicable, unknown) in adults

    c. All-cause childhood mortality (<5 years)
3. To map the distribution of AIDS (≥15 years) and all-cause mortality
4. To briefly describe the mortality profile of the population (≥15 years) with a particular focus on AIDS mortality

## 1.3 Literature Review

This section documents literatures on two areas: Firstly, the literature on the burden and impact of HIV/AIDS in South Africa and secondly, the literature on using GIS as a research tool in public health research both at the larger scale (national level) and local scale (district level). There is a growing body of literatures on the uses of GIS for health research from simple mapping of disease incidence/prevalence all the way to sophisticated risk models [8]. For instance, early records of disease maps date back to the first map made by Dr. John Snow in 1855 to trace the source of a cholera outbreak in London, England [8]. Whilst working in the Soho district of London at the time of a severe cholera epidemic, Snow [9] developed a spot map of Cholera deaths. On the map he showed that the distribution of deaths were concentrated around one pump located in Broad Street. He concluded that drinking water from the pump triggered the epidemic in the district and shut the pump down (the epidemic ceased soon after this).

The literature on using GIS technology in health research examines its applications, the limitations and the benefits it has offered in terms of analyzing and integrating spatial datasets. The use of GIS for health or disease research stretches across numerous disciplines including, biostatistics, epidemiology, public health, and other related disciplines [10]. However, although found to be under-utilized it has been concluded that GIS has much to contribute to health research in terms of strategic health planning and intervention [11].

## Definitions

The users of GIS have defined GIS in different ways, based on their functions, basic components, and uses in a way that suits them in their given working environment. For instance, Antenucci [12] defined GIS as a "computer system that stores and links nongraphic attributes or geographically referenced data with graphic map features to allow a wide range of information processing and display operations, as well as map production, analysis, and modeling." For the moment, however, the following definitions of the key words are deemed to be adequate for this study:

- *Geographic Information System (GIS):* Computer database that store and manipulate geographic data [13]
- *Spatial Analysis:* The study of the locations and shapes of geographic features and relationships between them [14]
- *SaTScan statistics*: Used for detecting and evaluate clusters in a temporal, spatial or space-time setting [7]
- *Spatial Clustering*: defined as a statistically significant excess or deficit of mortality events relative to expectation (Poisson model) as measured by the Kulldorff spatial scan [11]
- *AIDS Mortality*: defined as deaths that includes all deaths from AIDS with or without TB [55]

### 1.3.1 The HIV/AIDS burden in South Africa

In South Africa, HIV/AIDS are a major health concern. It has been reported in the UNAIDS/AIDS report that in 2005 alone, there were around 5.5 million living with HIV in South Africa, and almost 1,000 AIDS deaths have had occurred [41]. It is difficult to overstate the suffering that HIV has caused in South Africa. For each person living with HIV in South Africa, not only does it impact on their lives, but also those of their families, friends and wider communities. It is thought that almost half of all deaths in South Africa, and a staggering 71% of deaths among those aged between 15 and 49, are caused by AIDS [42]. So many people are dying from AIDS that in some parts of the country, cemeteries are running out of space for the dead [43]. A recent survey found that South Africans spent more time at funerals than they did having their hair cut, shopping or having barbecues [44].

A report published by the UNAIDS/WHO [41] states that, between 1990 and 2003 – a period during which HIV prevalence in South Africa increased dramatically – the country fell by 35 places in the Human Development Index, a global directory that ranks countries by how developed they are. Hospitals are struggling to cope with the number of HIV-related patients that they have to care for. In 2006 a leading researcher estimated that HIV-positive patients would soon account for 60-70% of medical expenditure in South African hospitals [45]. Schools have fewer teachers because of the AIDS epidemic. In 2006 it was estimated that 21% of teachers in South Africa were living with HIV [42].

It is very clear that AIDS is having a devastating impact on South Africa. There are many possible reasons why South Africa has been so badly affected by AIDS, including poverty, social instability and a lack of government action [33]. Based on literature review, AIDS is affecting mainly those who are sexually active, which means the demographics of the country are slowly changing. Many deaths are people who are their family's primary wage earners. This is resulting in many AIDS orphans who in many cases depend on the state for care and financial support.

### 1.3.2 HIV/AIDS and children

South Africa has never experienced an epidemic of the magnitude of HIV/AIDS. The routine surveillance conducted by the Department of Health has shown that among pregnant women attending public health clinics for antenatal care, the prevalence has increased from less than 1% in 1990 to 25% in 2001 [27]. The high prevalence of HIV/AIDS poses major challenges for both government and non-government organizations, who are doing their utmost to curb the spread of the disease and help those affected by it, by creating awareness around HIV/AIDS. With many women who are HIV-positive still not receiving drugs that could prevent them passing HIV to their babies, HIV infections are alarmingly common amongst children in South Africa.

According to government antenatal surveys, there were around 260,000 children aged below 15 living with HIV in South Africa in 2006 [46]. Children who are living with HIV are highly vulnerable to illness and death unless they are provided with pediatric antiretroviral treatment. Unfortunately, there is still a shortage of such treatment in South Africa. The AIDS Law Project, an NGO based in Johannesburg, estimated that 50,000 children in South Africa were in need of antiretroviral drugs at the beginning of 2006, but that only around 10,000 were receiving them [47]. As well as many children being infected with HIV in South Africa, many more are suffering from the loss of their parents and family members from AIDS. UNAIDS estimated that there were 1.2 million South African children orphaned by AIDS in 2005, compared to 780,000 in 2003 [41]. Once orphaned, these children are more likely to face poverty, poor health and a lack of access to education.

### 1.3.3 HIV/AIDS in KwaZulu- Natal (KZN)

It is evident that KwaZulu-Natal province of South Africa is widely categorized as the worst HIV/AIDS-affected region in the world. HIV prevalence rates within KZN have always been higher than any other province and approximately 10% higher than that national average, and this has been a trend that has largely been sustained since the first sero-prevalence testing of women attending public antenatal clinics was undertaken in 1990 [48]. According to Erskine [48], when the epidemic emerged in KZN, the areas of higher HIV prevalence were associated with major transport routes and areas of severe poverty.

Research has shown that as early as 1990 KZN had a HIV/AIDS prevalence twice that of the national level: 1.6 percent vs 0.76 percent [48]. In addition the data shows that those individuals who are under 25 years of age are more vulnerable to HIV/AIDS than other age groups. Apart from physical suffering, the pain of bereavement and the loss of household heads and breadwinners, it threatens to cut a swathe through the productive workforce, doing great damage to the economy [49]. Furthermore, the high proportion of young people in KZN is one of the causes for a higher-than-average HIV prevalence in the province.

### 1.3.4 HIV/AIDS in the Africa Centre HDSS

Looking particularly at the mortality pattern in the Africa Centre HDSS, it has been reported that HIV/AIDS has considerably increased mortality rates in the study population and significantly reduced life expectancy at birth [69], which implies that the population are at a higher risk of infection. A study done in 2000 showed that the probability of dying between the ages 16 and 60 years was estimated at 58% for women and 75% for men [69], respectively, and of these deaths AIDS was responsible for causing 73% and 61% of the female and male deaths, respectively, among the 15–44 age groups. HIV/AIDS is by far the leading cause of death, followed by non-communicable diseases [55]. However, a recent study has suggested that the upward trend in mortality rates is being reversed by the antiretroviral therapy (ART) programme which has contributed significantly to an increase in life expectancy [55,59].

The Africa Centre HDSS longitudinal data at the community level lays the platform to understand the impact of HIV/AIDS interventions at the individual and population level. This sort of data provide reliable, policy-relevant information on evolution of HIV epidemic in the area to describe its social and demographic consequences, and to measure coverage, uptake and impact of national prevention, treatment and care programmes. Information on causes of deaths in the area were obtained via the VA approach which was an effective way to find out what the actual causes of mortality are, and was deemed necessary as it is obvious that most of the people do not die in hospital, but at home. The data collected have been extensively used to provide empirical evidence about the demographic and social impact of the HIV epidemic in a severely affected population.

Understanding patterns of HIV/AIDS transmission within rural areas requires knowledge about patterns or circulation and about sexual contacts between residents and their non-resident partners. Circular migration has also been identified as an important determinant in the rapidly progressing HIV epidemic in rural areas of Africa [69]. It is evident that HIV/AIDS is a major burden especially in rural Africa with a massive development challenge facing rural communities today and new approaches are warranted to properly manage and control HIV/AIDS at the district level (community level). Since many diseases are related to location, GIS through cluster detection analysis can be an appropriate approach to identify critical AIDS mortality locations for public health effort.

### 1.3.5 The Potential for using GIS

Many scholars and researchers have supported the potential of this technology based on how GIS has been employed to meet their specific needs. Verhasselt [15] discussed the potential of GIS in medical geography. According to Verhasselt, GIS facilitates hypothesis-generating because of its capability to: (1) overlay and integrate spatial information, and (2) substantiate quantitative analyses in disease ecology and health care delivery through its capability to handle large amounts of data. Scholten and de Lepper [16] also supported the idea of using a GIS because of its potential value to improve public health and understand environmental risk, and also recognizing the capability of GIS in managing and analyzing health care data.

Hendrickx et al.(2001) states that the use of GIS to many public health problems offers potential in assisting decisions on allocations of resources, prioritization of control areas, and planning and management of field operations [17]. In support of the above statement, the potential application of this technology can immensely assist health officers to target high risk areas and prioritize areas that are in need of health services/intervention and also can significantly improve adherence to treatment in Africa and elsewhere. GIS technology is a tool of great inherent potential for health in Africa as health is largely determined by environmental factors (including the sociocultural and physical environment) which vary greatly in space [23]. It can enable decision makers to easily assess where populations are in relation to specific problems (disease and health events) and in relation to available resources (health and social services, water supply and education, etc).

Today, GIS technology is now becoming increasingly important epidemiological tool in infectious disease epidemiology and control [18]. When properly used, GIS can inform and educate; empower decision-making at all levels; help in planning and cost-effective actions, in predicting outcomes and continually monitor and analyze changes over time making it one of the modern technological research tool ever developed to assist in disease intervention and control programmes. However, despite its potential it is evident that GIS use within Africa is limited, for instance there’s little current research on HIV (incidence/prevalence) that has utilized this technology to its full potential. It is perceived that the potential application of this technology in South Africa can offer exciting potential in understanding many of the public health problems at the local, provincial and national level.

### 1.3.6 Applications of GIS

Researchers working on malaria projects in Africa have widely used GIS in the management and control of malaria [19-21], to measure the effects of access to malaria treatment [22] and to evaluate the effects of intervention strategies [23]. Apart from its wide usage in malaria management in Africa, GIS has also been applied to other public health problems, for instance, on tuberculosis (TB) to help improve health services delivery (treatment). For instance, Tanser and Wilkinson [24], applied GIS and GPS technology to document and quantify improved access to TB treatment through a community-based programme in Hlabisa, South Africa. In their work they were able to plot TB supervision points used by the district health system in 1991 (programme's first year) and 1996 (programme fully established), and quantified access by using GIS to measure the mean distance from each homestead in the district to hospital, clinics, community health workers (CHW) and volunteer supervisors. The above study is one of a few GIS studies ever done at the local level to improve access to treatment in rural South Africa, and this approach can be adapted and subsequently be used at the national level.

Other programmes making use of GIS applications include WHO Public Health Mapping programmes which initially started in 1991, to boost efforts to eradicate guinea worm, poliomyelitis, onchocerciasis, roll back malaria, and other diseases [8]. Since then, the use of GIS have been greatly simplified and expanded to meet the distinct data needs of several disease control initiatives, including programmes for the elimination of onchocerciasis, blinding trachoma, African trypanosomiasis and lymphatic filariasis, as well as the global initiative to eradicate poliomyelitis and roll back malaria [8]. Within the health research discipline some of the applications of GIS are; assessing resource allocations, determining geographic distributions of the disease, mapping people at risk, mapping disease transmission, analyzing spatial and longitudinal trends, stratifying risk factors, strengthening disease surveillance (district to national level), planning and targeting resources/interventions and monitoring disease and interventions over time.

Today one sees a continual increase in the power and ease of use of GIS, better integration and easier availability of related technologies, such as GPS and spatial analysis. These tools are also diverse ranging from GIS as a core and unifying technology to geospatial statistics. Regardless of the application of this technology for example in health research, the end results must translate into improving levels of human health.

#### 1.3.6.1 Spatial Analysis (Clustering)

A search through the literature on spatial analysis has shown that geographers and other research scientists have used spatial analysis techniques to study disease patterns and health care delivery systems [11]. Since then, the literature has grown, most notably in two directions. First, some of the techniques described in the review article have become more sophisticated. Second, as predicted by Gesler, GIS has been increasingly used in applying spatial analysis [25]. Indeed, GIS technology has fostered a revival in the spatial analysis of health and disease phenomena, often facilitating the rapid calculation of appropriate formulas and the display of results. GIS adds the dimension of spatial analysis to information technology by providing the interface between information and a map. This makes it easy to present information to key decision-makers more quickly, efficiently and more effectively in addressing problems. Hence, knowledge of the geographic distribution of HIV/AIDS can inform and improve public health practice and interventions.

Studies in other parts of Africa have documented significant space-time clustering of malaria. For instance, Snow et al [26] report a space-time clustering of severe childhood on the Coast Kenya with seasonal peaks in incidence of severe malaria comprising discrete mini-epidemics. Similar studies on microepidemiology of malaria are now underway in the Nouna study area [6] and the results are likely to help better understand the observed clustering of mortality in the area. Thus the identification of clusters can lead to interventions, even when etiology is uncertain. Therefore, investigating whether the distribution of adverse health outcomes in a population are either random or not should be an important primary objective before starting a programme for primary and secondary prevention of health events.

Clustering detection methods are classified as global, local and focused [64, 65]. Global and local spatial statistics are exploratory tools that allow for the finding and identification of clusters without a pre-determined hypothesis about cluster location, whereas focal spatial statistics allow for testing the hypothesis of whether a disease, its vectors or its hosts cluster around a suspected location. Focal statistics are one approach to deal with fine resolution health data, since clustering of vector-borne diseases can occur at very short distances, as observed for dengue fever cases and *Ae. aegypti* mosquitoes [66, 67], for LaCrosse encephalitis [68] and for malaria [69]. In this study we used the local spatial statistic referred to as Kulldorff spatial scan statistic throughout this paper to detect statistically significant mortality clusters in the Africa Centre HDSS as it is deemed appropriate for the purpose of this study.

#### 1.3.6.2 GIS applications in South Africa

The public health sector in South Africa has, for some time now, widely used GIS and researched on how it can be applied. The Ministry of Health and the Medical Research Council are well known users of this technology in this sector [28]. The Ministry of

Health has focused a lot on using GIS to capture Health Service Infrastructure whereas more of the analytical research in terms of accessibility to health services has been conducted by the Medical Research Council [28]. However, from literature search there are evidence of few published examples of the use of GIS in health systems research in Africa but there is an encouraging amount of work in progress. For instance, only a handful of published studies using GIS applications to study TB in South Africa exist [31], and more studies in this area using this empirical mapping approach is encouraged.

One study group used GIS to study inequalities in population per bed ratios and the implications of open access to the private and the formerly white hospital services in the province of KZN, South Africa [32]. In an urban area in the Western Cape Province, South Africa, researchers used GIS to map the TB cases [33] and analyze childhood TB in two urban communities of Cape Town, South Africa [34]. As seen in these studies GIS has played an important role in terms of TB control programs, service development and research, however, more work is needed to promote the applications of this technology across all levels of health research and non-government organization and other stakeholders and scale up the intervention programmes from district to national level in the country (application in HIV-above section). In the rural district of Hlabisa, South Africa, Tanser and Colleague [29] used GIS to explore the relationship between HIV prevalence and the proximity of homesteads to primary and to secondary roads. An exploration of the causes of HIV heterogeneity is important as it may inform control efforts [29]. This type of study is very useful in estimating ones distance from where he/she is located (geographically) to the nearest road system, as proximity to road networks could be one factor associated with HIV prevalence and GIS greatly help in this regard. This is one of the few studies done (as reported in literature) in South Africa employing GIS applications.

Consequently, in this study we employed the applications of GIS and Kulldorff's spatial scan statistic to help us better understand the AIDS mortality patterns both in space and time across the Africa Centre HDSS. In light of the magnitude of HIV/AIDS in the area, we have applied GIS to be used as an adaptive management tool to help specify the problem of HIV/AIDS risk, and to target high risk communities and prioritize management interventions. Whilst GIS cannot solve the HIV/AIDS pandemic, it should be seen as an information-gathering tool that can be employed in the mobilization and response to the disease burden. This will enable focused and direct intervention where most benefit will be gained. We envisage that GIS can act as a powerful evidence-based practice tool for control, management and planning of HIV/AIDS intervention in the Africa Centre HDSS, and thus may provide a model for usage in other sectors of South Africa.

### 1.3.7 Limitations regarding GIS use

While some scholars were optimistically urging use of GIS technology, others were advocating caution before jumping on the GIS bandwagon [10]. The following are the main concerns expressed by some of the authors and researchers: The first concern lies in the lack of a "universally accepted definition" [38]. This is due to the fact that many organizations now using GIS basically define GIS according to how it benefits them in their respective working environment; therefore, it is hard to come up with a standard definition. The second concern relates to the issues of the suitability of a GIS to answer certain research questions [39]. In some instances, using a GIS evokes more questions than it does answers, which of course supports hypothesis generation, but also possibly fails to answer the original questions [39]. A third problem resides in the data-intensive nature of a GIS. The requisite disease and health information may be inaccessible. The underlying reasons for inaccessible data include not only its expense but also the political nature of establishing data sets [38]. The fourth concern is that, the data might not be of appropriate geographic scale (aggregation) or temporal frames to support specific investigations in questions [25, 39].

#### 1.3.7.1 Factors limiting the potential of GIS in South Africa

There are certain factors that have limited the full potential use of GIS in health in Africa. According to Tanser [23] the major impediments to the growth of GIS in Africa are lack of suitable GIS data sets and training (human capacity). Lack of suitable GIS data is a problem that has faced GIS users for decades in both the developed and developing nations. Without adequate data, GIS is not very useful. The access to spatial data (which are fundamental to any GIS application) continues to be difficult and expensive [40]. The problem with limited data source is the one that is really slowing the process and advancement of GIS in South Africa, and as a result not so much has been published about GIS work in Africa. Data is the very core of any GIS application. What needs to be done now is to extensively collect data starting at the district level, but this would require collective efforts from the government and other stake holders to support financially in this exercise. The fact that GIS is a relatively new technology means that staffs should be well-trained and equipped with necessary skills and knowledge in GIS. But for them to gain such knowledge staff should be attending training so that they can be able to use this technology successfully. If GIS are to be useful and effective, then they must be introduced by local scientists who understand both the technological and the socio-economic context in which the systems are to operate. This will help them to properly explore this technology and apply it more meaningfully. Hence, capacity development of African staff should therefore be prioritized.

Very little work has put sufficient emphasis on accurately reveling what the spatial dimension of public health problems actually looks like at the community level, especially in the rural areas. Furthermore, limited work has attempted to demonstrate how resources can effectively be distributed to problem areas. It therefore makes it very difficult for decision-makers to accurately know where HIV/AIDS problems are high, moderate, or low in terms of their prevalence be it at the village, provincial and national level. This in turn makes it difficult for them to know how best to manage them, distribute resources which are allocated to address them and determine which problems should be given priority over others, in an objective or scientific manner.

The first part of this study, chapter one (introduction) provides a short overview (background) of the topic concerned, the aims and objective of the study as well as the literature review. This is followed by the methodology (chapter two) applied in the study, while the third chapter presents the results from the study. The fourth chapter provides the discussion based on the findings and recommendations, and finally the fifth chapter (conclusions) sums up the major findings of the study.

## CHAPTER 2.0: METHODOLOGY

This chapter describes the methods employed in this present study and is organized in the following manner: Firstly, we describe the study area and population, then we describe the study design and database used, and finally we describe the geographic and statistical analysis methods.

### 2.1) Study area and population size

The study area (Fig.2) referred to as the Africa Centre HDSS throughout this thesis is located near the market town of Mtubatuba in the Umkanyakude district of KwaZulu-Natal province, South Africa. It was established in 2000 and since then has followed vital events in a population of 87,000 including all deaths.

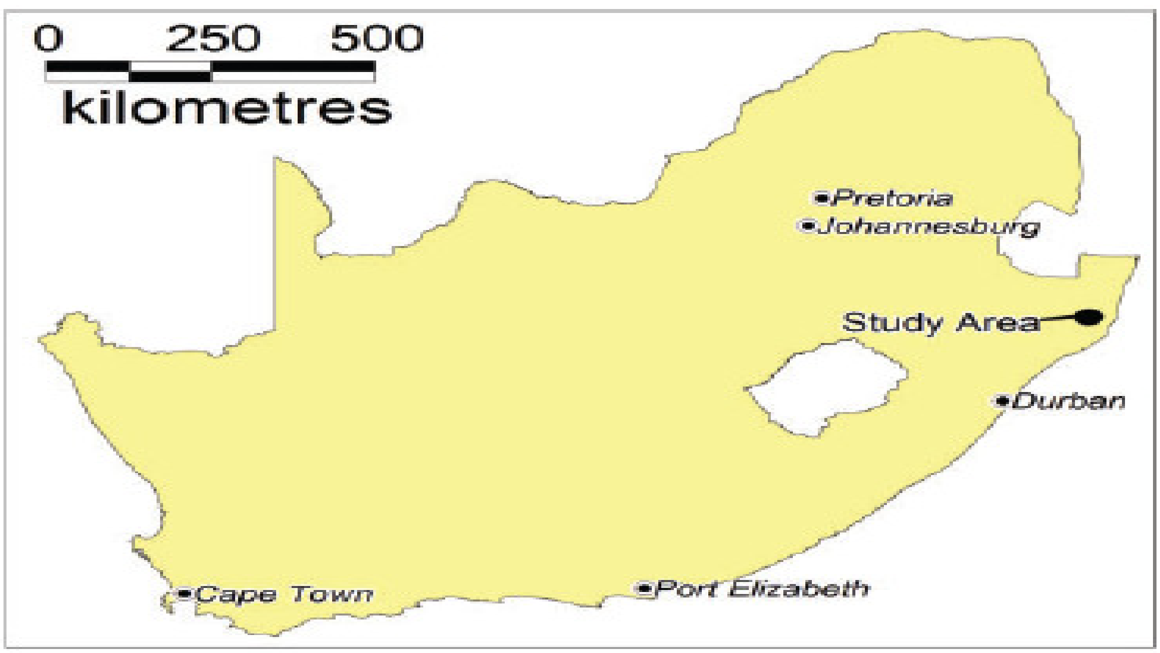


**Figure 2** Location of the study area in South Africa (Source: Tanser *et al*., 2007)

The entire study area has been mapped using a GIS system. All bounded structures (homesteads) within the HDSS have been mapped via global positioning system (GPS) technology and assigned numbers respectively. Mapping the entire area was useful in terms of tracking people who have moved from one homestead to another and also by equitably allocating field workers workload. The Africa Centre HDSS is a typical of many rural areas of South Africa in that while predominantly rural, it contains an urban township and informal peri-urban settlements [51]. The area is characterized by large variations in population densities (20–3000 people/km$^2$). Mostly in rural areas homesteads (households) are scattered rather than grouped. Furthermore, despite being a predominantly rural area, the principle source of income for most households is waged employment and state pensions rather than agriculture. Tanser and colleague [51] documented that, approximately 77 percent of households in the area had access to piped water and toilet facilities.

The population is made up of both resident and non-resident members of household. The non-residents are especially those who may be the head of household, bread winners or members of the family that are working elsewhere in other centres or townships, but still a full member of a household. Individuals can be members of more than one household at any point in time (e.g. polygamously married men whose wives maintain separate households). The surveillance area is 438 km$^2$ in size and includes a population of approximately 85, 000 people who are members of approximately 11, 000 households.

**2.2) Study design and database**

This is a secondary data analysis study using data from the Africa Centre HDSS population-based longitudinal cohort study. Residents mainly adult population (≥15 years) both males and females forms the study sample; this includes individuals who are member of a household within the surveillance area but not necessarily resident within it.

The datasets for this study comes from the main ACDIS database. This database stores data on registered subjects (individuals), physical structures (e.g. homesteads, clinics and schools) and household [55]. The information is updated every four months through fieldworker visits. The interviewers who collected the information on the cause of deaths are well trained according to standard standardized procedures. The events that are recorded include: individual events (e.g. death, migration, etc) household events (household formation, migration, change of household head), and events affecting bounded structures (start of a new building, change of a building's main purpose or its owner). The data collected are more reliable as events are recorded soon after they occur and can be validated in subsequent rounds.

We focused mainly on the AIDS-related mortality (deaths) events as our main outcome health event in the Africa Centre HDSS, therefore we only used the following available datasets from the main ACDIS database: data on causes of deaths (e.g. AIDS and all-cause), exposure data (person years) and GIS datasets (GPS points, base map and other vector datasets). The cause of death data and exposure data were geo-coded or linked to the GPS homestead locations. Thus the data from the ACDIS main database have been collected since 2000, and many efforts have been made to provide as complete and accurate a database as possible. For example, each recorded death was subject to a VA in which trained fieldworkers conducted interview with the caregivers of those who died to identify the cause of death in the community. The use of VA was an effective way to find out the actual causes of deaths and was deemed necessary as it is obvious that most of the people do not die in hospitals, but at home.

Information on the cause of death is necessary for rational public health planning, yet mortality data in South Africa is limited [52] and can be a problem. For instance, unreliable census data undermine the population base necessary for calculation of rates [53]. In the Africa Centre HDSS causes of deaths are grouped into five main categories: AIDS, injuries, communicable, non-communicable and unknown [55]. AIDS is of course a communicable disease; however due to its magnitude as a cause of mortality, it is considered separately. The table below (table 1) gives the description of the valuables and causes of deaths. All the data were inputted using a MS-SQL Server database [51].

**Table 1** Mortality variable list

| **Variable** | **Description** |
|---|---|
| IIntID | Individual Internal ID (Individuals' refer to the individual members of households within the HDSS) |
| BSIntID | Bounded Structure Internal ID (house or building) |
| ExpYear | Year of exposure (observation) |
| AgeGrp | Numerical 5 year age group , 0=0-4, 1=5-9, etc |
| sAgeGrp | String age group |
| Sex | Male or Female member |
| ExpDays | Days of exposure (observation) for this individual in this year, age group and BSIntID |
| Died | 1 if the person died at the end of this residency episode (person could have died as a non resident later on) |
| Longitude | X coordinates of the bounded structure (BS) captured via GPS |
| Latitude | Y coordinates of the bounded structure (BS) captured via GPS |
| **Cause of deaths (COD)** | |
| Injuries | All cause due to intentional (homicide) and unintentional injuries (eg.road accidents) |
| Non-communicable disease | Eg. deaths from cerebrovascular disease, heart disease or neoplasm's |
| Communicable disease | Eg. Disease of the pre-transitional period, infectious, maternal, perinatal and nutritional-related cause of disease |
| AIDS | It is of course a communicable disease; however due to its magnitude as a cause of mortality in ACDIS, it is considered separately this group includes all deaths from AIDS with or without tuberculosis |
| Unknown or ill-defined | Insufficient information or no consensus about cause of death |

## 2.3 Geographic and statistical analysis

### 2.3.1 Mortality rates for the ACDIS data

Longitudinal HDSS provide a solid platform to understand the demographic and health research transition. The Africa Centre HDSS uses the VA approach to establish the major causes of deaths in which trained fieldworkers conducted interview with the caregivers of those who died to identify the main cause of death in the community [54]. The VA approach is most accurate for those causes of death associated with a set of signs and symptoms that easily distinguish it from other causes of death and that can easily be recalled and recounted to the interviewer [55]. However, many common causes of death present with signs and symptoms are not easily distinguishable.

We used mortality event data and person-exposure data (extracted from the main ACDIS database) to describe the mortality patterns of HIV/AIDS with TB for adult population among the Africa Centre HDSS population. When calculating the death rates (DR) we used the summed person time to calculate the crude rate by year and sex for the cause of deaths for different categories over the seven year period (2000-2006). We also used the summed person to calculate the adjusted rates (age-sex). By using the summed person time we were able to get an estimate of the actual time-at risk in years, months, or days that all persons (>15 years) contributed over the seven year period. The data were aggregated by local area within MapInfo program version 7.0, and exported to excel spreadsheet where rates were calculated and plotted on graphs.

We also computed standardized mortality rate for the adult population to eliminate the influence of age and sex composition using the direct standardization method. We applied the standard standardization rates set by the INDEPTH (an international network of health and demographic surveillance systems in Africa, Asia and Oceania) as a reference and calculated the rates accordingly. The standardization of mortality rates across the Africa Centre HDSS allows for comparisons as the population structures may differ at the community level. For example, since the highest age-specific mortality rates occur at the youngest and oldest age cohorts, populations with large child and elderly populations will have higher mortality rates so adjusting for age and sex will control for variations in adult mortality thus giving a more precise estimation of the mortality rate across the study area. Analyses were performed for adults across the two time periods (2000-2003 and 2004-2006), respectively. Maps of the mortality distribution rates were produced in MapInfo.

### 2.3.2 Kulldorff's spatial scan statistic (cluster detection method)

There exist several methods for disease cluster analysis that has been suggested by others [6, 70], however, we used Kulldorff's spatial scan statistic [7] in which the spatial distribution of the population is taken into consideration. This was done by gradually scanning a window across time and/or space across the study area, noting the number of observed and expected observations inside the window at each location using a Poisson model [7, 35, 36]. Thus, other statistical methods such as Kaplan Meier could have also been used, however, it is deemed appropriate to use the Poisson method to estimate the mortality rates so that the spatial estimate tally with the mortality rates. For any given position of the centre, the radius of the circle changes continuously so that it can take any value from zero up to a specified maximum value. For each circle, the spatial scan statistic calculates the likelihood of observing the observed number of cases inside and outside the circle [7, 71]. The circle with the maximum likelihood is the most likely cluster (least likely to have occurred by chance). A homestead is captured if it lies in the circle. The method creates a set containing an infinite number of distinct circles [6, 7, 37]. Each of these circles could contain a different set of neighboring homesteads and each of these circles is a potential cluster of mortality. Under the null hypothesis of complete spatial randomness, simulated maps of homesteads mortality cases conditioned on the total number of observed cases were repeatedly generated (999 times), and likelihood ratio test statistics were calculated for each potential cluster [35, 7]. The sets of potential clusters were then rank ordered according to the magnitude of their likelihood ratio test statistic.

Let $N$ be the total number of deaths in the study area, $n$ the observed number of deaths within the circle, and $\lambda$ the expected number of deaths in the circle under the null hypothesis. Let the number of deaths in each homestead follow a Poisson distribution.
Hence the likelihood ratio for a specific circle is therefore proportional to

$$\frac{L_A^{(D)}}{L_0} \propto \left[\frac{n}{\lambda}\right]^{n} \left(\frac{N-n}{N-\lambda}\right)^{N-n} I_{(n>\lambda)} \qquad (1)$$

Where $L_A^{(D)}$ is the likelihood under the alternative hypothesis that there is a cluster of elevated annual mortality rates in aged 15 years and above in a specific circle $D$, $L_0$ is the likelihood under the null hypothesis, and $I$ is an indicator function that is equal to $I$ when the circle has more deaths than expected under the null hypothesis, and 0 otherwise. Maximizing (1) over all circles results in the one that constitutes the most likely mortality cluster. The test statistic is

$$\max_{D} \frac{L_A(D)}{L_0} \qquad (2)$$

Kulldorff (8) has derived the likelihood ratio test and provided the properties of the test statistic. We used SaTScan 7.0.3 to perform the spatial analysis scanning to detect communities representing potential clusters (high or low rates) across Africa Centre HDSS. The spatial scan statistic implemented in SaTScan, was used to test for the presence of mortality spatial clusters and to identify their approximate locations. Spatial scan method [35] offers several advantages: it corrects for multiple comparisons, adjusts for the heterogeneous population densities among the different areas in the study, detects and identifies the location of the clusters without prior specification of their suspected location or size thereby overcoming pre-selection bias, and the method allows for adjustment for covariates [36, 37]. The *p*-value of the statistic is obtained through Monte Carlo hypothesis testing, where the null hypothesis of no cluster is rejected at an α level of 0.05 exactly if the simulated *p*-values is ≤ 0.05 for the most likely cluster [71, 7]. We ran the analysis by two time periods 2000-2003 (before ART programme) and 2004-2006 (after ART program) for AIDS, all-cause and injuries for adult population using purely spatial analysis scanning for high or low rates. We also ran the analysis for the whole period (2000-2006) for deaths due to unknown causes and all cause mortality (<5 years). We limited any possible cluster so it would not exceed 50% of the total Africa Centre HDSS population at risk.

### 2.3.3 GIS and Cartographic display of mortality clusters

All GIS manipulation and cartographic displays were performed in MapInfo which was used as both database and a visualization tool. MapInfo transforms data into information simply by integrating different datasets (mortality data, GPS coordinates and other spatial datasets), data querying and providing an output all in the manner designed for easy visualization and interpretation of the results. The mortality datasets and the homestead location data (GPS points) were imported into MapInfo for data aggregation. This was achieved by geocoding the mortality datasets to the GPS coordinates. Individuals were geo-located by linking them (using homestead points) to their last recorded homestead position [51]. Mortality events and person-exposure were aggregated by homestead using MapInfo. We used MapInfo to create maps (unstandardized and standardized age-sex ≥15 years) for time periods 2000-2003 and 2004-2006 for AIDS and all-cause mortality, respectively.

Although general GIS visualization tools and proposed methods for visualizing surveillance data exist [56] there isn't any clearly defined tools specifically tailored to perform SaTScan data visualization. Consequently, creating SaTScan visualizations requires parsing, cleaning and formatting the data for use in one of many GIS applications [56, 71]. In this study the calculated results of potential clusters in SaTScan where converted to DataBase File format (dbf.) and were imported into MapInfo to generate maps showing statistically significant mortality clusters (core areas) in the Africa Centre HDSS. Data visualization tools can aid analysts who must interpret complex data sets, because humans must ultimately determine which events require further investigation. We employ this empirical mapping approach to analyze the spatial dynamics of AIDS-related mortality patterns, thereby exploring the potential for the application of GIS to HIV/AIDS management. The produced disease maps would be very helpful for understanding small-scale variation in HIV/AIDS distribution.

## CHAPTER 3.0: RESULTS

The focus of this chapter is on the results of the study undertaken. The majority of the results are presented on maps with inserted data, as well as on graphs and table. The results are presented in the following order: Firstly, results on general death patterns, followed by all-cause, AIDS and other categories of mortality, respectively.

### 3.1 Africa Centre HDSS population profile

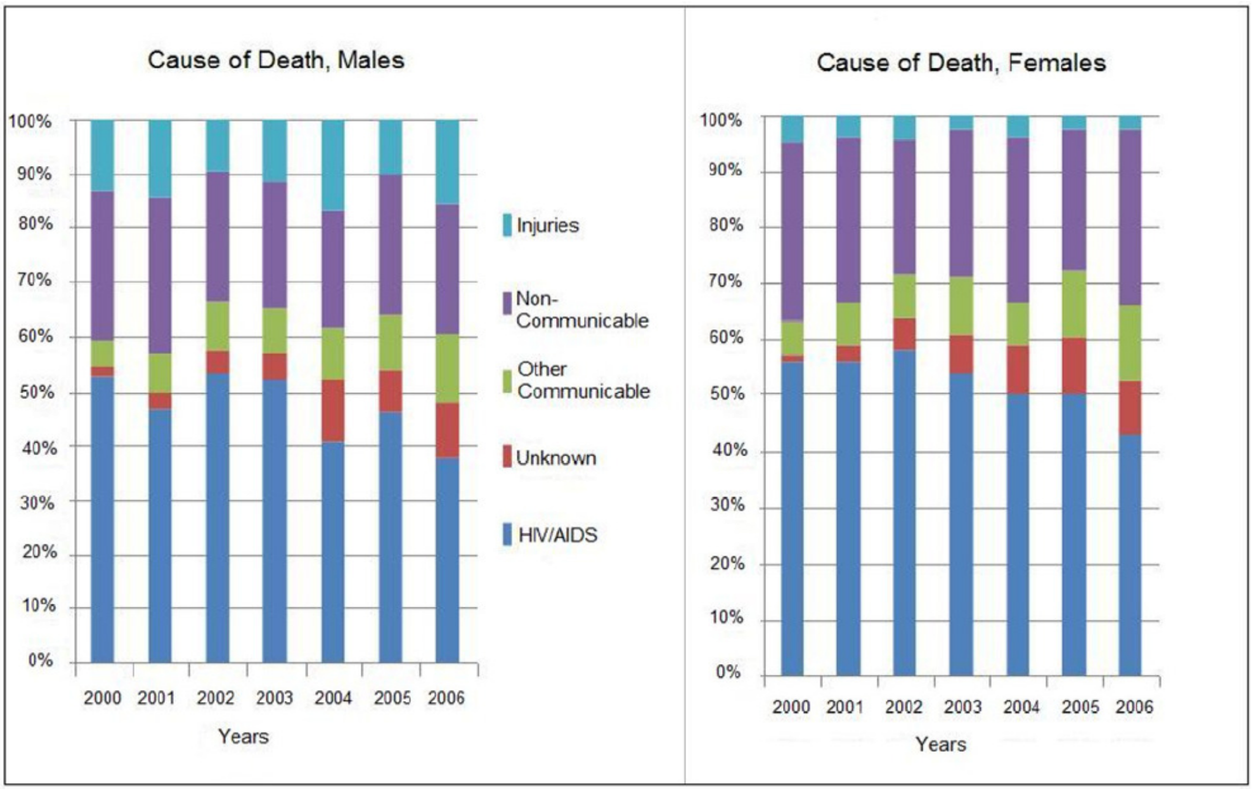


**Figure 3** Causes of deaths by year and sex for adult population during 2000-2006

Figure 3 presents the different categories of causes of deaths in the HDSS for adult population (males and females). The results show that HIV/AIDS with TB remains the major cause of death for both males and females (adult population) during 2000-2006. The second major cause of death is non-communicable diseases followed by other communicable diseases. Among males injuries is also high and also a significant cause of death. HIV/AIDS account for over 45% for males during the peak of the epidemic and for females it account for over 50%. The burden of HIV/AIDS and deaths is higher for females as compared to males.

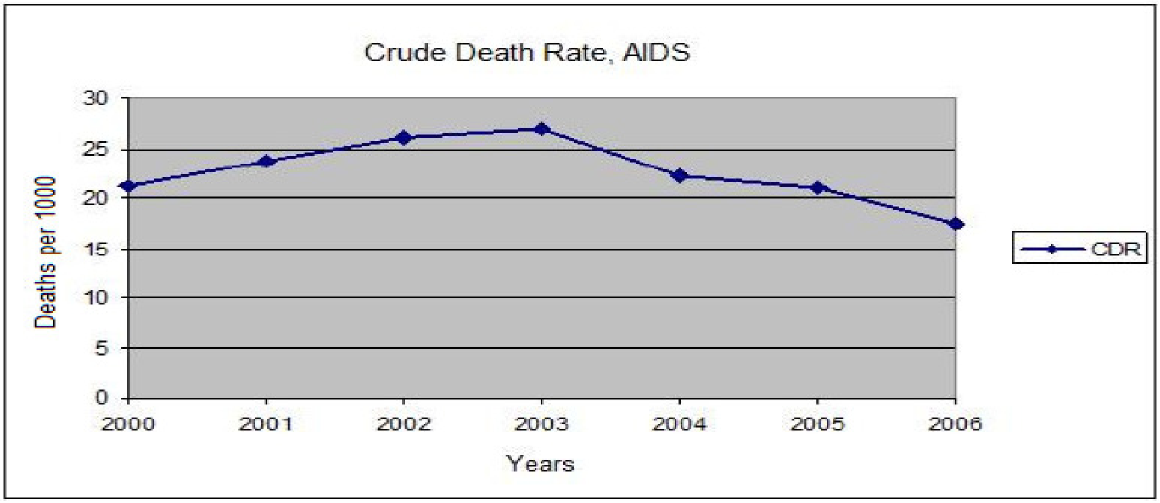


**Figure 4** Crude death rate for the 7 year period for both males and females (adult population)

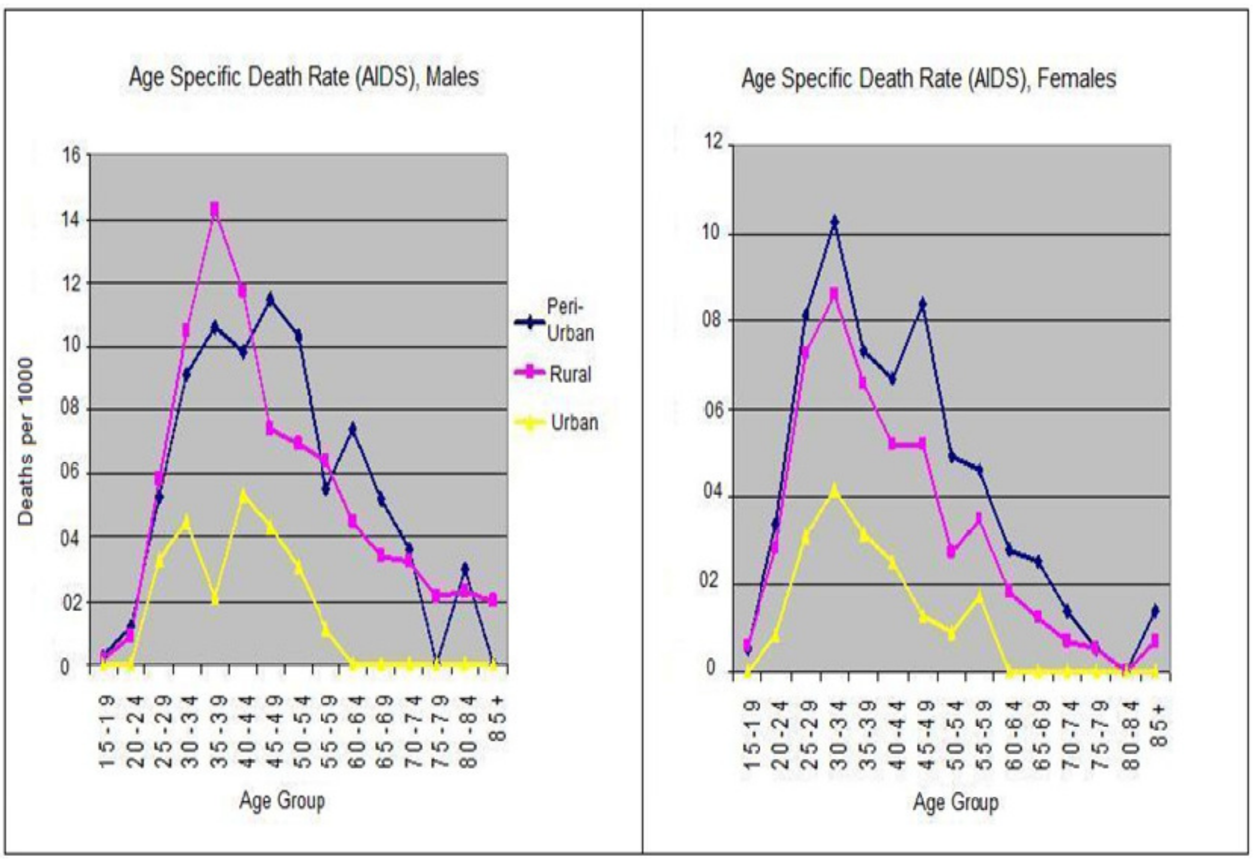


**Figure 5** Age specific death rates for males and females (≥15 years) by rural_urban settings, 2000-2006

Figure 4 shows the AIDS crude death rate (CDR) for the seven year period for both male and female (adult population) aged 15 years and above. The average yearly CDR (per 1000) for adult population was 22.6 for the seven year period. There is an increase in death rate starting from year 2000 through to 2003, after 2003 there is a decline (Fig.4).

Figure 5 presents age specific death rates for adult population stratified by settlement type (urban, peri-urban, and rural). In the rural areas there is evidence of high mortality rates (AIDS) for males (15-39 years) and for females (15-34 years) high mortality rates found especially in the peri-urban areas (Fig.5). Figure 5 also suggests a higher death rate for males as compared to females.

### 3.2 Spatial clustering of all-cause mortality

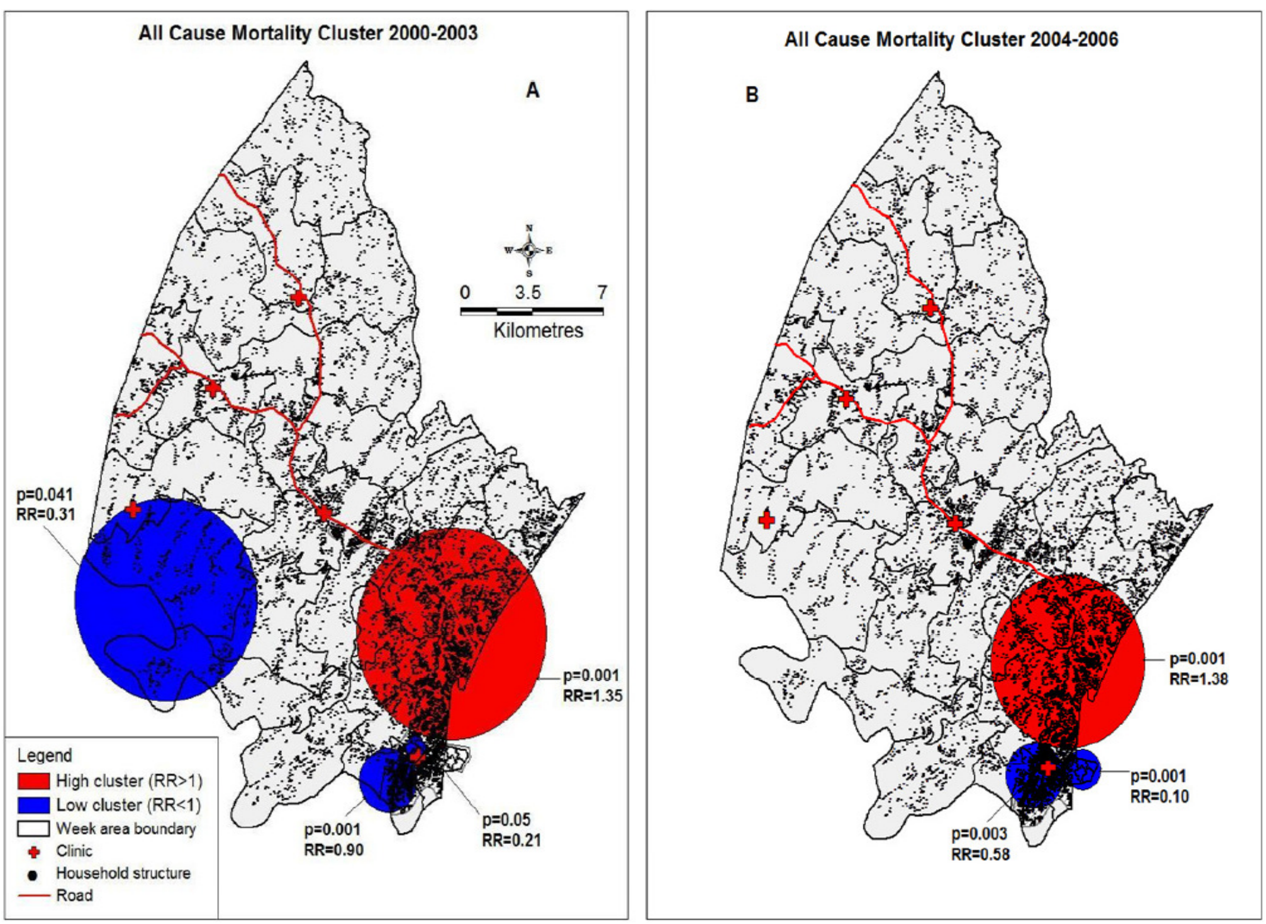


**Figure 6** All-cause mortality cluster for time periods 2000-2003 and 2004-2006 (≥15 years)

Figure 6 presents the results of the purely spatial analysis scanning for high or low rates of all-cause mortality cases for adult population for time periods 2000-2003 (Map A) and 2004-2006 (Map B). Kulldorff's spatial scan statistic has identified statistically significant clusters with high and low relative risk for both time periods. For time period 2000-2003 a statistically significant cluster (*p*=0.001) was detected with a high relative risk of 1.35: 1194 total observed cases (973.66 expected), whilst for time period 2004-2006 a statistically significant cluster (*p*=0.001) was detected with a high relative risk of 1.38: 660 total observed cases (518.07 expected).

### 3.3 Spatial clustering of AIDS mortality

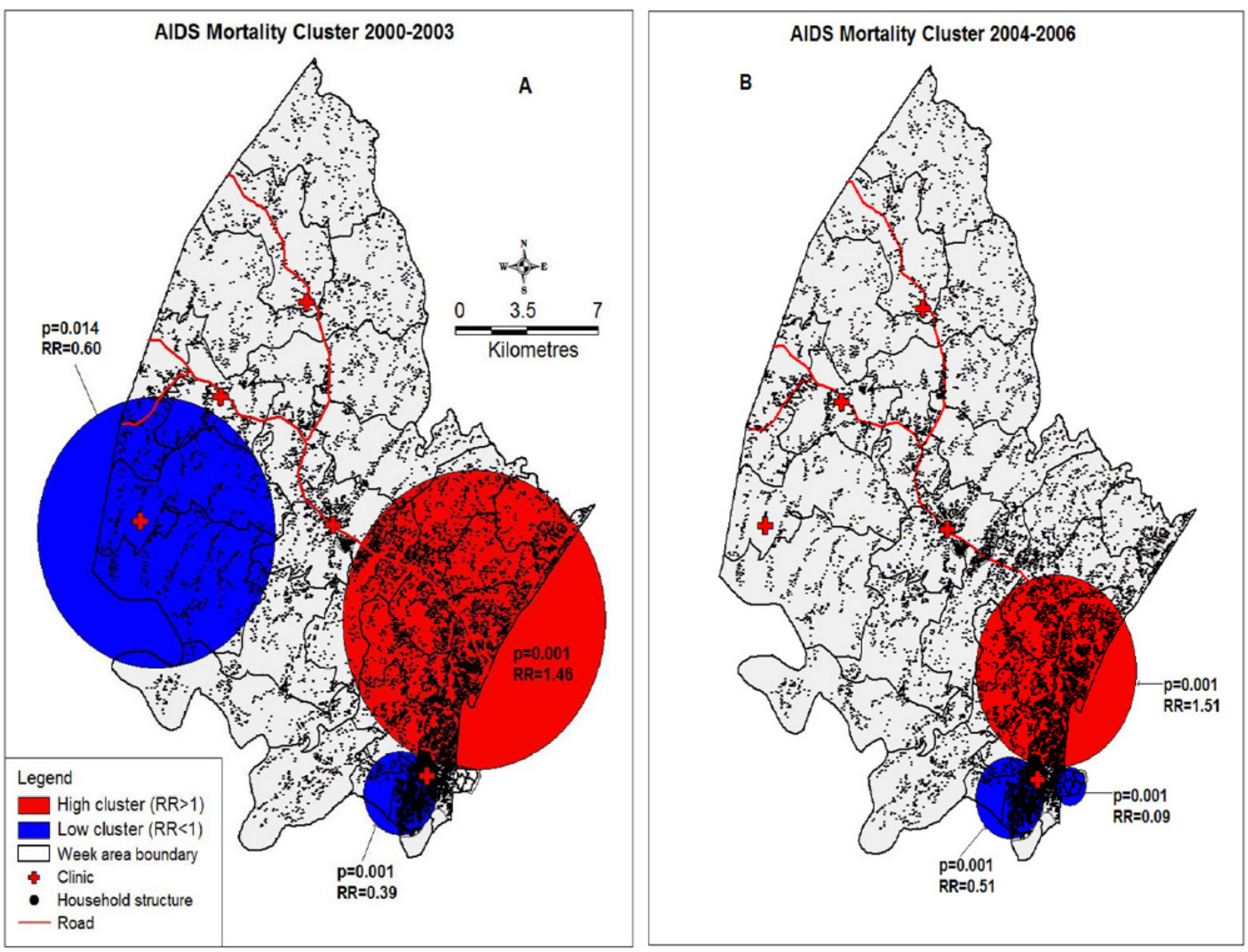

**Figure 7** AIDS mortality clusters for time period 2000-03 and 2004-06 for adult population

Figure 7 presents the results of the purely spatial analysis scanning for high or low rates of AIDS (with TB) mortality cases for adult population for time periods 2000-2003 (Map A) and 2004-2006 (Map B) as detected by the Kulldorff's spatial scan statistic method. For time period 2000-2003 a statistically significant cluster ($p$=0.001) with a high relative risk (RR) of 1.46 was detected: 836 observed AIDS mortality cases (673.24 expected). For time period 2004-2006, a statistically significant cluster ($p$=0.001) with a high relative risk (1.52) was detected: 323 observed cases (237.30 expected).

Thus the potential high clusters with a high relative risk of 1.46 (Fig.7, Map A) and 1.51 (Fig.7 Map B) have also confirmed that the areas detected were known to have high HIV/AIDS prevalence among the residents. A consistent low-risk cluster was detected at around KwaMsane township across both time periods (RR=0.39, 0.51). The result shows reduction in mortality for time period 2004-2006 as evident by the decrease in the size of significant high rate cluster as shown on Figure.7, Map B.

## 3.4 All-cause and AIDS mortality rates

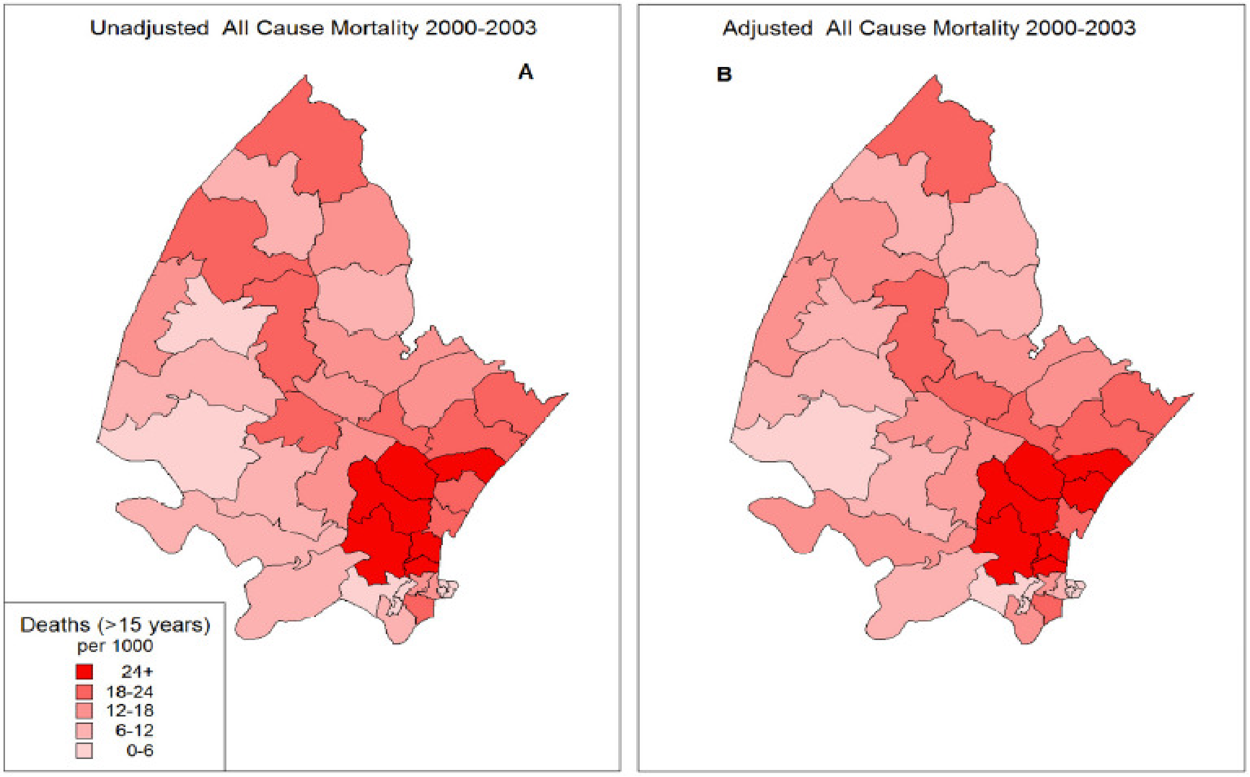


**Figure 8** Crude and standardized rates of all-cause mortality for adult population, 2000-2003

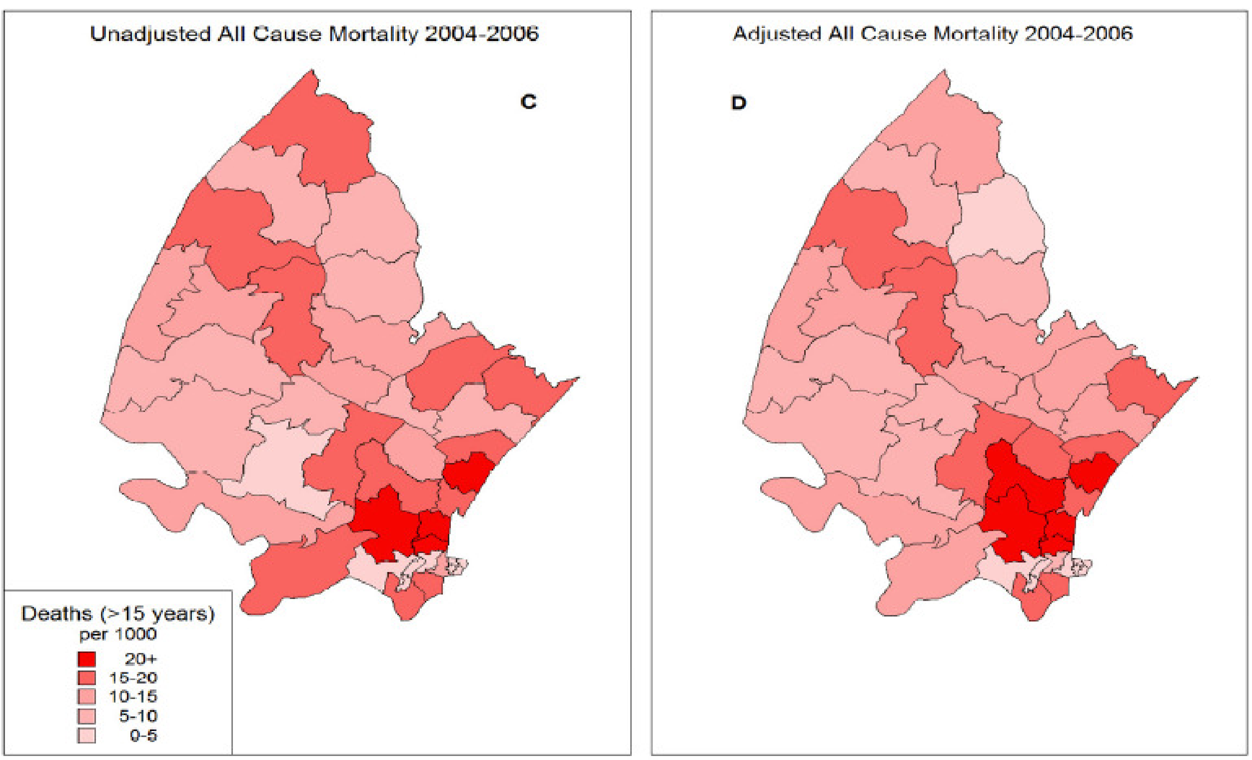


**Figure 9** Crude and standardized rates of all-cause mortality for adult population, 2004-2006

Figure 8 shows the crude (Map A) and standardized (Map B) rates of all-cause mortality for adult population across the HDSS for time period 2000-2003, whilst figure 9 presents the crude (Map C) and standardized (Map D) rates of all-cause mortality for time period 2004-2006. All-cause mortality results (Fig.8 and 9) have shown similar pattern as shown below by AIDS mortality distribution rate (Fig.10 and 11). It is seen that for time period 2004-2006 (all-cause mortality) there is a decrease in general mortality rates.

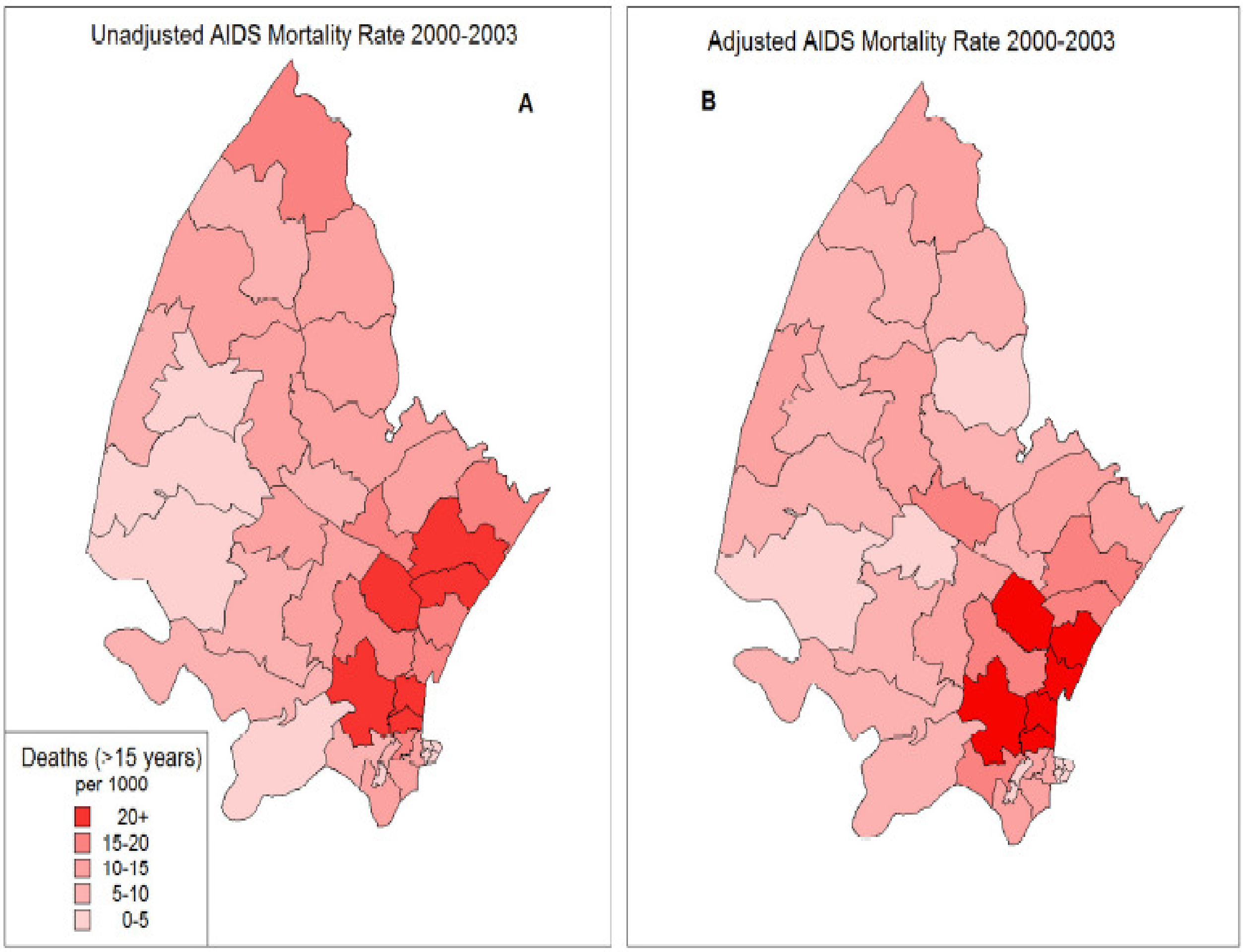


**Figure 10** Crude and standardized (age-sex) rates of AIDS mortality for adult population, 2000-2003

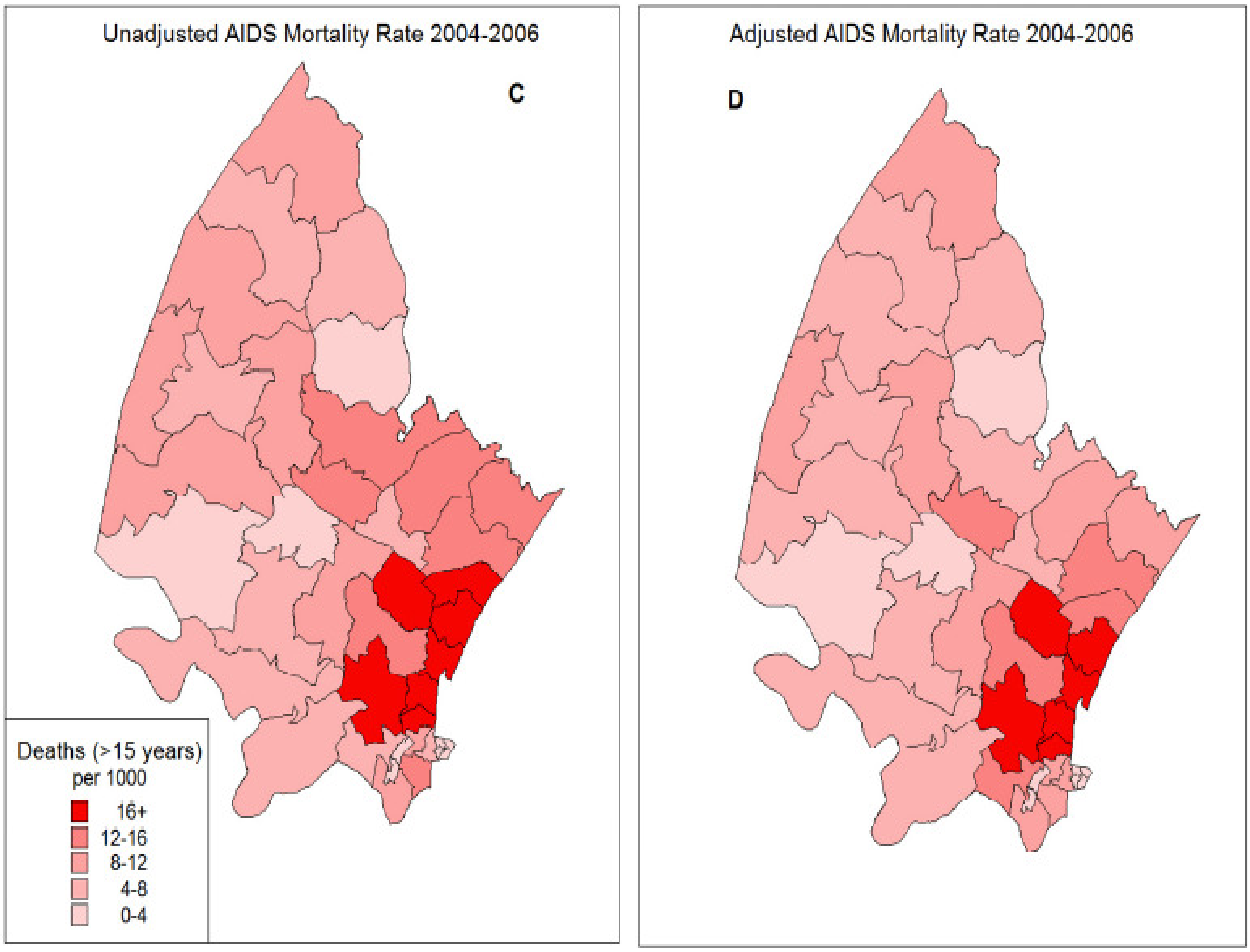


**Figure 11** Crude and standardized (age-sex) rates of AIDS mortality for adult population, 2004-2006

Figure 10 shows crude (Map A) and standardized rates (Map B) of AIDS mortality for adult population across the HDSS for time period 2000-2003. Death rate 20 and above indicates areas or communities with higher AIDS mortality rate. Figure 11 also presents the crude (Map C) and standardized (Map D) rates of AIDS mortality distribution rates but for time period 2004-2006. It is seen that for time period 2004-2006 there is a decrease in AIDS mortality in the HDSS as compared to time period 2000-2003 (Fig. 10). There were no major differences in mortality distribution rates as observed on the two maps (crude and standardized) for time periods 2000-2003 and 2004-2006.

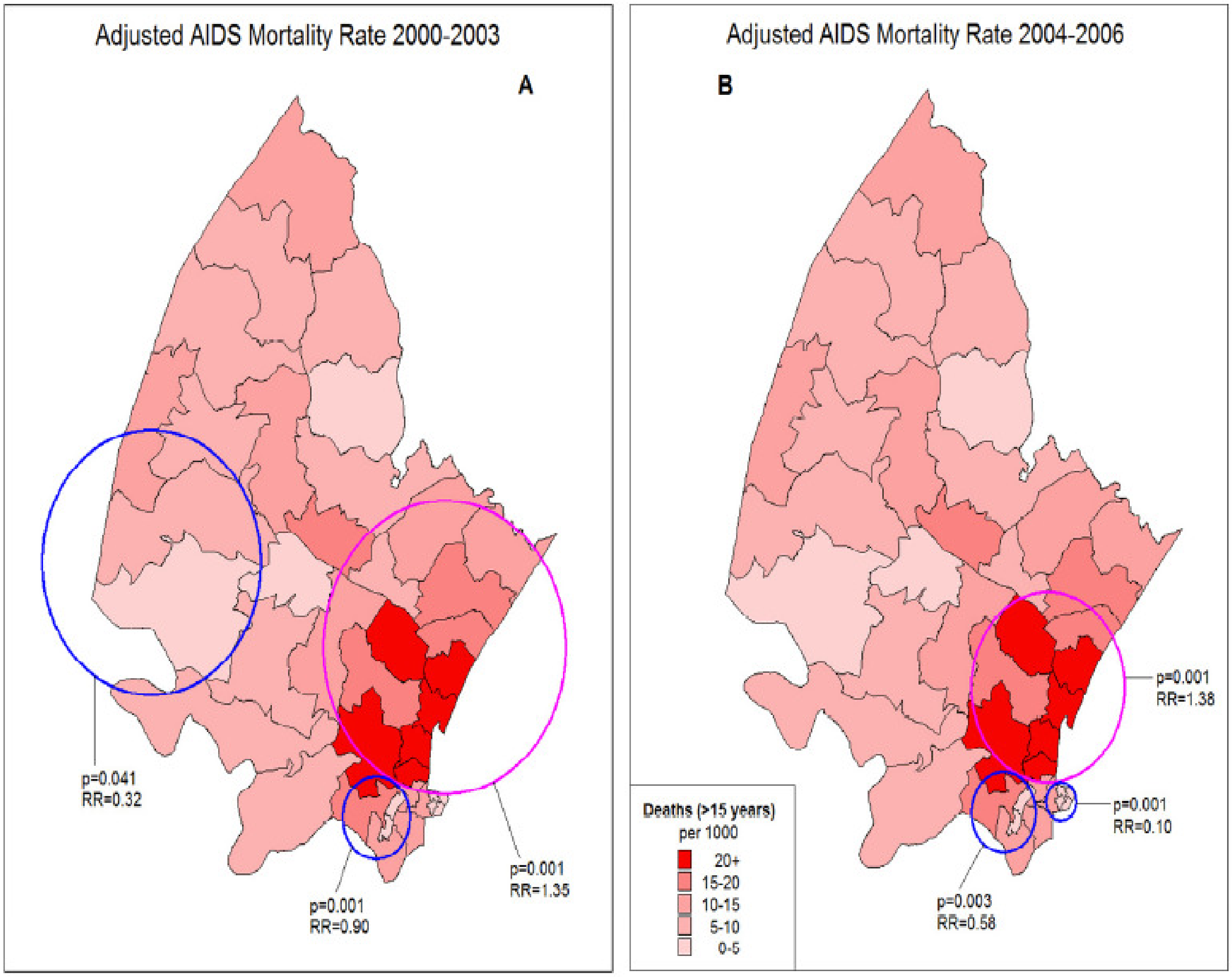

**Figure 12** AIDS standardized maps across the two time periods 2000-2003 and 2004-2006

Figure 12 presents the AIDS standardized rates for the two time periods 2000-2003 and 2004-2006 for adult population across the HDSS (the cluster results shown above is from Fig.7 with same p-values and RR). The statistically significant clusters were superimposed onto the standardized maps (base map) for comparison purposes for instance, to compare the significant cluster results with the standardized mortality distribution rates to geographically identify high and low risk areas in the community. The above result reveals that for the two time periods significant clusters with a high relative risk were detected around communities (areas) where there is existence of higher mortality rates.

## 3.5 Spatial clustering of other cause of deaths categories

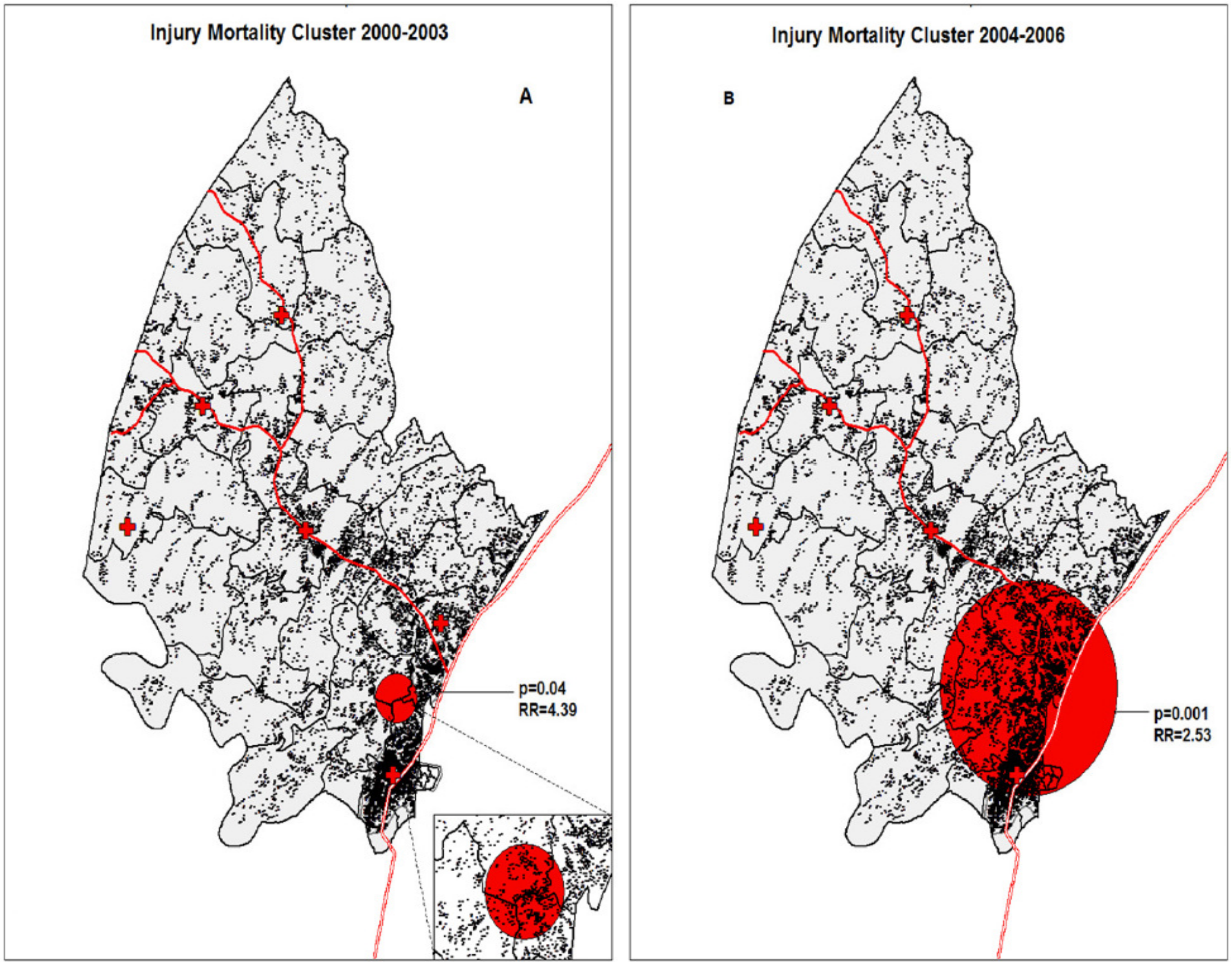


**Figure 13** Injury mortality cluster for time periods 2000-2003 and 2004-2006 for adult population

Figure 13 shows the identified statistically significant clusters of injury mortality as measured by the Kulldorff's spatial clustering method for both time periods. Deaths attributable to injuries also exhibit strong spatial clustering tendencies ($p \leq 0.05$) with a high relative risk of 4.39 and 2.53). However, there were no identified statistically significant clusters with low relative risk (rates) detected for both time periods as measured by the Kulldorff's method.

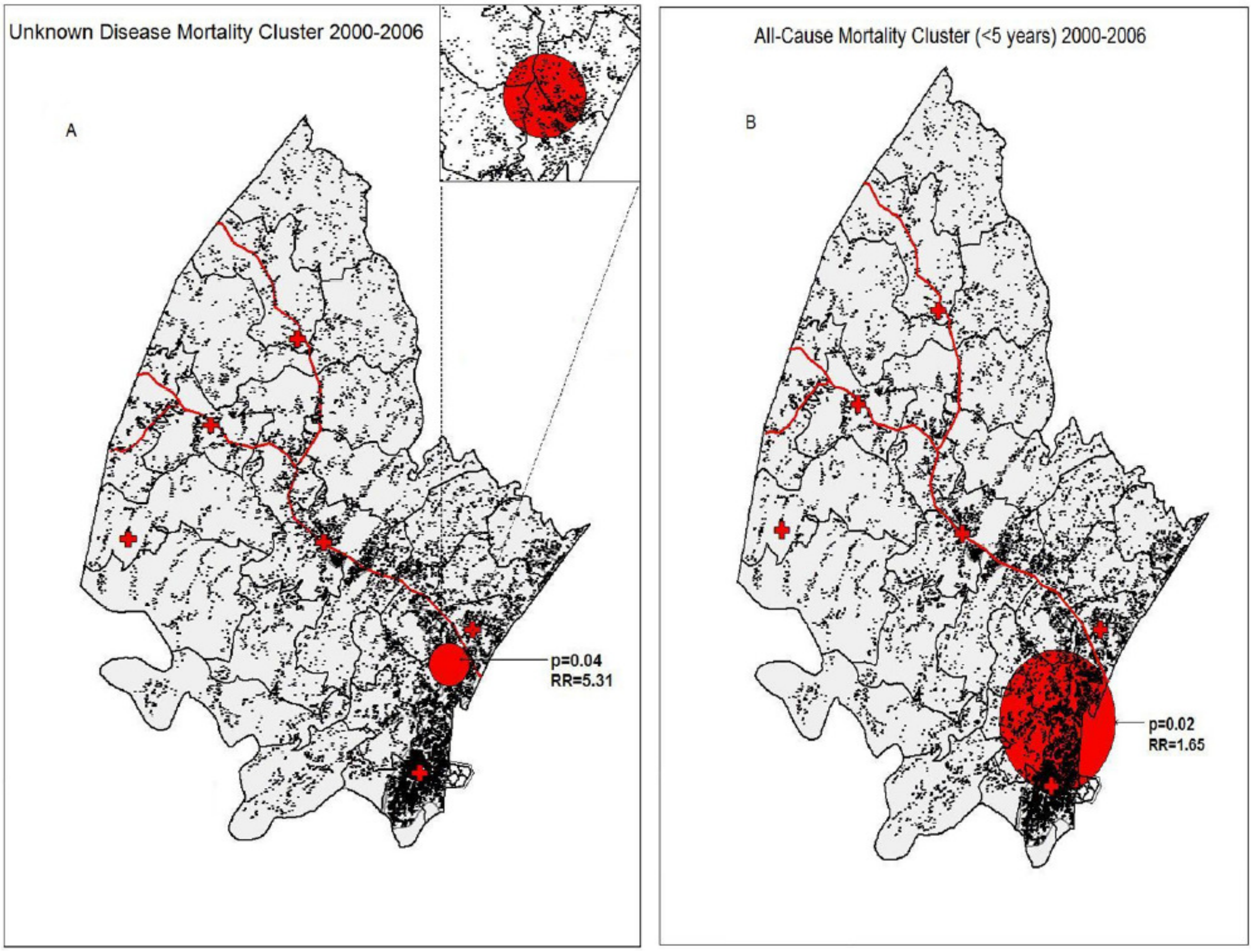

**Figure 14** Unknown and all-cause childhood mortality (<5 years) cluster for the whole period, 2000-2006

Figure 14 presents the results obtained from the purely spatial analysis scanning for high or low rates of morality cases attributable to unknown causes (Map A) and all-cause childhood mortality (Map B) for the whole period 2000-2006. Unknown and all-cause childhood mortality also exhibit strong spatial clustering tendencies with a high relative risks of 5.31: 14 observed cases (2.72 expected cases) for unknown mortality cases whilst for all-cause childhood mortality a high relative risk of 1.65: 134 total observed cases (85.50 expected cases) were observed for the whole period (2000-2006). No statistically significant clusters with low rates were detected for unknown and all-cause childhood mortality for the whole period.

**Table 2** Summary table of other communicable and non-communicable mortality clusters

| **Cause of death** | **Time period** | **Cluster type** | **Observed Cases** | **Expected cases** | **Relative Risk (RR)** | ***p* value** |
|---|---|---|---|---|---|---|
| Other Communicable disease | 2000-2003 | Most likely | 92 | 63.78 | 1.68 | 0.69 |
| Other Communicable disease | 2004-2006 | Most likely | 9 | 1.26 | 7.25 | 0.29 |
| Non-communicable disease | 2000-2003 | Most likely | 341 | 281.05 | 1.35 | 0.51 |
| Non-communicable disease | 2004-2006 | Most likely | 30 | 62.97 | 0.45 | 0.07 |

Table 2 shows the summary results of other communicable and non-communicable mortality cases obtained using purely spatial analysis scanning for high or low rates for time periods 2000-2003 and 2004-2006. After running the analysis for the two time periods for other communicable and non-communicable mortality it is seen that there were no identified statistically significant clusters as measured by the Kulldorff's spatial scan statistic. Despite no detected statistically significant cluster the result suggest a high relative risk of 1.68 for other communicable (2000-2003) and 7.25 (2004-2006), whilst for non-communicable a high relative risk of 1.35 (2000-2003) and a low relative risk of 0.45 (2004-2006) were detected.

## CHAPTER 4.0: DISCUSSION

In this chapter we firstly discuss the results of the study undertaken based on the findings, secondly, we discuss the chosen empirical mapping approach and its limitations and finally, on the recommendations for management approach.

### 4.1 Africa Centre HDSS Population

In sub-Saharan Africa there is little reliable national-level data describing the distribution of deaths by cause, and the WHO's mortality database [78] lists HIV-related causes for only one sub-Saharan nation: South Africa in 1996 when roughly 4% of all deaths were attributed to HIV-related causes. Given the lack of reliable national-level data on cause of death, the South African study instead relied on indirect methods to estimate the levels of HIV-related mortality, and it is among very few credible national-level studies of HIV-related mortality in sub-Saharan Africa [79]. However, to fill this gap regarding limited data there is now available longitudinal data from Africa Centre HDSS at the local scale (district level) on cause of death to understand the impact of HIV related deaths at the individual and population level and thus this data is one of those current, up-to-date and most reliable sources of data available.

It is evident that HIV/AIDS (Fig.3) is the leading cause of death for both males and females (adulthood), however, the burden is much higher for females because of higher female HIV prevalence. Young adults, especially women, have been particularly affected by HIV/AIDS. Statistics South Africa reports that between 1997 and 2004 the death rate for women age 20-39 more than tripled and for men 30-44 more than doubled. The death rate peaks at 30-34 for women and 35-39 for men [77]. This parallels UNAIDS's estimate that women make up 58% percent the adult population (15 and over) who are HIV-positive in South Africa. Similarly, in Magu HDSS in Tanzania, a study was done that used the community based HIV cohort study to examine trends in HIV prevalence and incidence since 1994. In this particular study the result showed that HIV prevalence remains higher in women than man and higher in the trading centre than in rural areas [75]. Since 2000 in Africa Centre HDSS AIDS morality rate increases till 2003 (peak year of the epidemic) and afterwards decreases (Fig.4). It appears that more deaths (AIDS) occur in rural areas for males whilst for females in peri-urban areas (Fig.5). The peak of dying for males is between the ages of 15-39 years whilst for females 15-34 years. A report on adult mortality (age 15-64) between 1997 and 2004 by Statistics South Africa, based on death certificates, shows a dramatic increase in the death rate, largely due to AIDS-related deaths [77]. These are potentially young and active age groups who are or maybe the ones responsible for generating incomes (bread winners) for their respective families.

The potential of the Africa Centre HDSS as a research and health intervention platform was set up in 2000 as a HIV surveillance site, primarily to describe the demographic, social and health impact of the HIV epidemic in a population going through the health transition and to monitor and evaluate the impact of the intervention strategies on the epidemic [50]. The platform offers an opportunity to implement a wide range of studies that aim at improving the health of the population by producing evidence for policy, including testing new health products. The platform also provide relevant information to researchers, scientists, institutions, INDEPTH students and others in the planning and organizing of various research activities at the Africa Centre HDSS. Since its inception several research studies and implementation programmes have been carried out. The data used are fine in a sense that the data are current, constantly updated over time (annually since 2000), and most reliable providing an up-to-date information in a truly longitudinal manner at the community level [52]. Ethical clearance for this study has been obtained from the University of KwaZulu-Natal and was approved in terms of a secondary data analysis study.

### 4.2 AIDS mortality clustering

A purely spatial analysis scanning for high or low rates of AIDS cases were performed for time periods 2000-2003 and 2004-2006 for an adult population thus using a Poisson model. AIDS (Fig.10) exhibit strong spatial clustering tendencies as measured by the Kulldorff's method. The low-risk clusters of RR=0.39 (2000-2003) and RR=0.51 (2004-2006) detected around KwaMsane township for AIDS mortality across both time periods could be a result of early and sustained access to antiretrovirals (ARVs) in that particular community. Other countries show the same trend, for instance, a study in Tanzania show the similar pattern in AIDS mortality decline following an ART programme rollout in 2004. A significant decline of AIDS mortality was observed after the introduction of the ART rollout [76]. It is seen that communities that lies within the high cluster are the areas with high mortality rates. This is further confirmed by comparing the standardized rates (Fig.12) for the two time periods with the mortality clustering results. Interestingly, excess mortality clusters always occurred in the East of the study area along the National Road centred mainly around the Nkodibe and Bhobosa local areas thus further work is required to look into potential risk factors responsible for these excess mortality clusters in this particular location. One ethical difficulty with spatial analysis is identifying "hot spots” or specific locations (eg. Ingweni villages) thus bringing into question confidentiality issue. However, our approach in this study does not aggregate by any unit and instead the data define the "communities" and circles typically overlap several locations at once making stigmatization unlikely.

In late 1990 as reported by Hosegood and colleagues, HIV/AIDS mortality in the area rose sharply, and since then this population has experienced a sudden and massive rise in adult mortality [57]. This can be accounted for by the high number of AIDS deaths among adulthood. There is, however, preliminary evidence to indicate that mortality in general is on the decline in the area. The decline in mortality in the area has been attributed to the likely impact of ART programme, rolled out on large scale in the late 2004 [58]. The decline in mortality after 2003 in the Africa Centre HDSS is further confirmed in this study as shown in figures 4, 7 and 11. However, further work (socio-economic, health seeking behaviour) is needed to understand the underlying mechanisms responsible for the spatial clustering. For instance, a study in East, Central and Southern Africa argued that in low-income sub-Saharan Africans countries, where poverty is widespread, increasing access to resources for women may initially increase risk of HIV [74]. Communities located within the low clusters are located close to health facilities (clinics) and this could also have an effect as having easy access (proximity) to health facilities and services improves ones health.

For comparison against AIDS mortality patterns we also performed the analysis for all-cause mortality (≥15 and <5 years) and other causes of deaths (injuries, other communicable, non-communicable and unknown). All-cause mortality includes all categories of causes of deaths. All-cause mortality exhibit strong spatial clustering tendencies for the two time periods (Fig.6). It is seen that AIDS is driving all-cause mortality patterns which is thus influenced by the high number of AIDS mortality reported cases compared to other causes of deaths. The low cluster (RR=0.1) identified at Ingweni villages (Fig.6, Map B) is likely an artifact of the extension of the HDSS in 2004 but deserves further investigation. Analysis were performed across the two time periods for other communicable and non-communicable diseases, however, no statistically significant clusters were detected (Table 2). This is likely due to less number of reported mortality cases for these two causes of deaths. Injuries, unknown and all-cause childhood mortality were all detected as statistically significant clusters for each respective time periods (2000-2003, 2004-2006 and 2000-2006).

Injuries, mostly resulting from road traffic accidents or violence are a significant cause of deaths for adult males (15-44 years). For unknown causes further investigation is needed to establish these deaths as there is evidence of a high RR (5.31). All-cause childhood mortality also exhibit strong spatial clustering tendencies for the whole period (2000-2006), however, no statistically significant cluster with a low RR was detected. The significant high cluster result for all-cause childhood (Fig.14, Map B) could be influenced by AIDS-related mortality event so further investigation is also warranted to identify the exact underlying cause of death.

### 4.3 Empirical mapping approach

Mapping is a powerful tool for helping people visualize the geographic distribution of health problems as well as the structural and ecologic context of these problems and GIS is an innovative technology for generating this type of information. In developing countries GIS is being applied more often to answer complex problems and inform optimal resource allocation. In the Hlabisa subdistrict in rural South Africa, GIS has been used to document the effect of community-based TB treatment on distance to treatment location [24], to equitably distribute fieldworker workload in a large population-based survey [72], and to investigate an ecological relationship between proximity to roads and HIV prevalence among women attending prenatal clinics [29]. Therefore, using GIS and longitudinal surveillance data it can be a powerful tool for monitoring and management of disease and other public health programmes. Recent evidence suggests that geographic approaches to control and prevention may enhance public health efforts [63] and improve public health care delivery systems. Although GIS output does commonly appear as a geographical presentation, the tool is far more powerful than its basic mapping abilities—it is better described as a spatial analytical system [73].

The last two decades have seen substantial work in the GIS arena on the robust statistical analysis of point patterns, e.g, Besag and Newell (1991) [63], and thus spatial clustering is one of the many statistical analysis used for point pattern analysis. Spatial clustering detection methods are classified as global, local and focused [63, 64]. Global and local spatial statistics are exploratory tools that allow for the finding and identification of clusters without a pre-determined hypothesis about cluster location, whereas focal spatial statistics allow for testing the hypothesis of whether a disease, its vectors or its hosts cluster around a suspected location. We chose the local spatial statistic known as the Kulldorff's spatial statistic (used throughout this thesis) over the others [60, 61] as it is suitable for the purpose of this study because it determines the existence of statistically significant clusters and their geographic locations. The SaTScan software used for spatial clustering is freely available and can be downloaded via the internet on the following link: http://www.satscan.org. The software is user friendly and comes with a user manual which can be downloaded. The results from the analysis in SaTScan can then be imported into any GIS software applications (e.g. MapInfo or others) for visualization and interpretation. Although the MapInfo software is user friendly, one still needs to be taught how to use it especially beginners thus training is recommended. We have used the methodology on retrospective data, however, it can also be applied on prospective data.

The development and evaluation of effective programmes to reduce the burden of disease requires a detailed knowledge of the spatial distribution of the disease or mortality and causal pathways [6] since many diseases are related to location. Quite often public health authorities need to respond to demands to investigate potential clusters of different diseases and confirm or refute, with certainty, whether a problem exists [71]. Such an approach makes explicit the rationale for targeting treatment interventions and contributes to the development of community as well as political support for health interventions [63]. Consequently, in this study we used the chosen mapping approach as a quick research tool to investigate significant mortality clusters in the Africa Centre HDSS. The benefits of this study can lead to proper planning of intervention programmes as HIV/AIDS is the most important public health problem in the area and interventions are clearly warranted (same approach can be applied to other diseases as well). The result of this study can help health officers to target high risk communities and prioritize the areas that need thorough epidemiological investigation and carry out intervention. The methods explored in this present study can subsequently be applied to larger scale data and be used to produce maps for future control programmes.

### 4.4 Limitations

Data on mortality collected through interview/survey is based on VA and thus the VA approach has an inherent limitation as it depends on recall of events and symptoms by relatives and absence of medical records and this can lead to or can be subjected to misclassification errors. The unknown causes of deaths may also influence the results or can lead the analysis to have a wrong conclusion, since many unknown or ill-defined cases are not fully resolved to determine the actual cause of death. Moreover, due to sample size limitations for unknown and all-cause childhood mortality; we only ran the analysis for the whole period and not for the two time periods. The MapInfo software used is not freely available like any other GIS software's which are freely available on the internet, therefore, the end users need to purchase a license to operate this software.

A limitation of the analysis using the Kulldorff's spatial analysis scanning method is that clusters are defined as circles [6, 14] and therefore needs to be interpreted with caution. For instance, if a homestead with a lower mortality rate is located next to or is surrounded by homesteads with a higher mortality rates it is more likely that the homestead with a lower mortality rate will always be included in the cluster [30, 31]. This is most likely due to the fact that the borders of the clusters are uncertain and therefore need to be interpreted with caution. The uncertainty of the borders arises because often there are many windows that overlap with the potential cluster. These windows may have only slightly lower mortality rates and therefore it is probable that some of these areas get erroneously included in the cluster [62, 59]. It is also possible that some secondary clusters in the neighborhood of the primary clusters might actually be parts of the primary cluster [36].

### 4.5 Recommendations

It is very clear that HIV/AIDS is a leading cause of deaths and a major burden in rural South Africa, and thus requires a different response and policy decision, and so information must be available that reflects a realistic assessment of the situation at the district, provincial and national level. This study may be regarded as a first step in prioritizing areas that needs further investigation and public health intervention. With this sort of analytical study health officers can come up with better strategic health management plans for intervention. Thus the below points are recommended in the hope that health officers and decision makers may consider and more accurately determine areas where they can target intervention by determining:

- Where people living with HIV/AIDS reside in the Africa Centre HDSS in terms of their village, community, local/district municipality location,
- The areas where there appears to be high/low risk areas,
- The socio-economic group which is mainly infected and the relationship that exists between the groups and the infection rate, and
- In which areas resources are mainly needed, which resources and how much e.g. condoms, marketing campaigns, education, drugs (ARVs), services etc.

## CHAPTER 5.0: CONCLUSION

It is a forgone conclusion that we live in a society where the resources that we use to address the multitude of problems we face are limited. It is imperative therefore that we explore all the possibilities information technology can offer us in enabling us to use our resources in an efficient and sustainable manner. As an information technology, WHO describe GIS as, "an excellent means of analyzing epidemiological data, revealing trends, dependencies and inter-relationships that would be more difficult to discover" using traditional tabular approaches [8]. Moreover, today the applications of GIS are becoming increasingly important in infectious disease epidemiology and control and thus with disease maps health officers can come up with better management approaches and focus on areas that need intervention the most.

The results suggest the existence of high and low rates of statistically significant clusters of HIV/AIDS mortality. This study extends our understanding of the geographic distribution of mortality at the local level (community) and may aid our understanding of significant causes of deaths among the adult population. This study has also demonstrated that using existing health data, GIS and spatial scan statistics could provide public health officials with additional tools necessary for disease surveillance at targeting of control measures in a resource limited settings. More detailed individual level epidemiological investigations are needed to be carried out in the identified clusters of significantly high mortality to identify the most important determinants of disease distribution and assess the burden of illness due to mortality in the Africa Centre HDSS.

## 6.0 REFERENCES


1. Merson, M.H. The HIV-AIDS pandemic at 25--the global response. N Engl J Med 2006. 354: p. 2414-2417
2. CDC. The Global HIV/AIDS Pandemic. MMWR, 2006. 55(31): p. 841-844.
3. WHO. The World Health Report 2000. Health Systems: Improving performance. World Health Organisation: Geneva.
4. WHO. The World Health Report 2004: Geneva.
5. Centre for Actuarial Research. The Demographic Impact of HIV/AIDS in South Africa. National and Provincial Indicators for 2006. 2006, South African Medical Research Council and Actuarial Society of South Africa.
6. Sankoh, O.A., Yazoume, Y., Saueborn R, Muller O & Becher H. Clustering of childhood mortality in rural Burkina Faso. Int. J. Epi, 2001.30: p.485-492.
7. Kulldorff, M. A spatial scan statistic. Communications in Statistics: Theory and Methods, 1997. 26: p. 1481-1496.
8. WHO. Public Health Mapping 2003: Geneva.
9. Snow, J. On the Mode of Communication of Cholera. 1855, Second Edition. J. Churchill, London, 1855. Reprinted in Snow on Cholera, The Commonwealth Fund, New York,1936.
10. Albert, D.P., Albert,W., Gesler, M., Levergood, B., Fellers, R.A. & Messina, J.P. Spatial Analysis, GIS, and Remote Sensing Applications in the Health Sciences. 2000, Ann Arbor Press. p. 1-9.
11. Gesler, W. The uses of spatial analysis in medical geography: a review. Soc Sci Med 1986. 23: p. 963-73.
12. Antenucci, J.C . Geographic Information systems: A Guide to the Technology. 1991, Van Nostrand Reinhold, New York.
13. Aronoff, S. Geographic information systems:A management Perspective. Ottawa: WDL Publications, 1989.
14. Environmental Systems Research Institute. Introduction to ArcView GIS. 1996, Two day Course Notebook with Exercise and training Data.RedLands,California: Environmental Systems Research ,Inc.
15. Verhasselt, Y. Geography of Health: Some trends and perspectives. Soc Sci and Med, 1993. 36(2): p. 119-123.
16. Scholten, H.J., & Lepper, M.J.D. The benefits of the applications of geographic information systems in public and environmental health. World Health Statistics Quarterly, 1991. 44(3): p. 160-170.
17. Hendrickx, G., de La Rocque, S., Reid, R., Wint, W. Spatial trypanosomosis managment:fom data-layer to decision making. Trends in Parasitology, 2001. 17(1): p. 35-41.
18. Michael, E. & Bundy, D.A.P. Global Mapping of Lymphatic Filariasis. Parasitology Today, 1997. 13(12): p. 472-476.
19. Hightower, A.W., Ombok, M., Otieno, R., Odhiambo, R., Oloo, A.J, Lal, A., Nahlen L & Hawley W.A. A geographic information system applied to a malaria field study in western Kenya. Am J Trop Med Hyg,1998. 58:266-72.

20. Martin, C., Curtis, B., Fraser, C. & Sharp, B. The use of a GIS-based malaria information system for malaria research and control in South Africa. Health Place, 2002. 8: p. 227-36.
21. Booman, M., Durrheim, D.N., La Grange, K., Martin, C., Mabuza, A.M., Zitha, A., Mbokazi, F.M., Fraser, C. & Sharp, B.L. Using a geographical information system to plan a malaria control programme in South Africa. Bull World Health Organ 2000. 78: p. 1438-44
22. Schellenberg, J.A., Newell, J.N., Snow, R.W., Mung'ala, V., Marsh, K., Smith, P.G. & Hayes, R.J. An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya. Int J Epidemiol,1998.27:323-9.
23. Thomson, M.C., Connor, S.J, D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M. & Greenwood, B. Predicting malaria infection in Gambian children from satellite data and bed net use surveys:the importance of spatial correlation in the interpretation of results. Am J Trop Med Hyg, 1999. 61.2-8.
24. Wilkinson,D., Tanser, F. GIS/GPS to document increased access to community-based treatment for tuberculosis in Africa. THE LANCET, 1999. 354: 394-5.
25. Albert, D.P., Gesler, W.M. & Wittie, P.S. Geographic information systems and health: An educational resource. Journal of Geography,1995.94(2):350-356.
26. Snow, R., Armstrong-Schellenberg, J.R.M., Peshu, N et al. Periodicity and space-time clustering of severe childhood malaria on the coast of Kenya Trans R Soc Trop Med Hyg, 1993. 87: p. 386-90.
27. Department of Health 2002. National HIV and Syphilis Antenatal Sero-prevalence Survey in South Africa, 2002-2006, http://www.doh.gov.za.
28. Tengbeh, S.M. & Martin, S.M. The Application of Geographical Information Systems (GIS) in Public Health Management: A Case of HIV/AIDS: www.gissa.org.za. Volume,
29. Tanser, F., LeSueur, D., Solarsh, G. & David Wilkinson. HIV heterogeneity and proximity of homestead to roads in rural South Africa: an exploration using a geographical information system. Tropical Medicine and International Health, 2000. 5(1): p. 40–46.
30. UNAIDS 1998. AIDS epidemic update: Geneva.
31. Tanser, C.F. & Le Sueur, D. The application of geographical information systems to important public health problems in Africa. Int J of Health Geographics, 2002. 1(4).
32. Zwarenstein, M., Krige, D. & Wolff, B. The use of a geographical information system for hospital catchment area research in Natal/KwaZulu. S Afr Med J 1991. 80: p. 497-500.
33. Beyers, N., Gie R.P., Zietsman, H.L., Kunneke, M., Hauman, J., Tatley, M. & Donald, P.R. The use of a geographical information system (GIS) to evaluate the distribution of tuberculosis in a high-incidence community. S Afr Med J, 1996. 86: p. 40-1.
34. Van Rie, A., Beyers, N., Gie R.P., Kunneke, M., Zietsman, L. & Donald, P.R. Childhood tuberculosis in an urban population in South Africa:burden and risk factor. Arch Dis Child 1999. 80: p. 433-7.

35. Kulldorff, M., Rand, K., Gherman, G., Williams, G., DeFrancesco, D. Software for the spatial and space-time scan statistics. SaTScan. and M. Version 2.1. National Cancer Institute Bethesda, USA; 1998.
36. Kulldorff, M., Nagarwalla, N. Spatial disease clusters: detection and inference. Stat Med, 1995. 14: p. 799-810.
37. Kulldorff, M. Prospective time periodic geographical disease surveillance using a scan statistic. J R Stat Soc Ser A 2001, 164:61-72.
38. Taylor, P.J. & Overton, M. Further thoughts on geography and GIS. Environmental and planning, 1991. 23: p. 1087-1094.
39. Heywood, I. Geographic information systems in social sciences. Environmental and planning, 1990. 22(7): p. 849-854.
40. Briggs, D.J. & E.P. The use of geographical information systems in studies on environment and health. World Health Stat Q 1995. 48: p. 85-94.
41. UNAIDS/WHO. UNAIDS 2006 Report on the global AIDS epidemic. http://www.unaids.org.
42. Centre for Actuarial Research (2006). The Demographic Impact of HIV/AIDS in South Africa - National and Provincial Indicators for 2006.
43. The New York Times (2004) South Africa recycles graves for AIDS victims: http://www.nytimes.com.
44. SA Advertising Research Foundation 2004. All Media Products Survey: http://www.saarf.co.za.
45. Inter Press Service News Agency 2006. Health South Africa: a burden that will only become heavier. http://www.ipsnews.net.
46. Department of Health (2002-2006). National HIV and Syphilis Antenatal Sero-prevalence Survey in South Africa: http://www.doh.gov.za
47. Mail and Guardian Online (2006). ARV Programme Less than the Sum of its (Monetary): http://www.mg.co.za.
48. Erskine, S (2005). Health Economics and HIV/AIDS Research Division. www.heard.org.za.
49. Smith. HIV/AIDS in Kwa-Zulu Natal and South Africa, AIDS Analysis Africa, Southern Africa Edition. 2000. 11(1).
50. Solarsh, G., Benzler J, Hosegood V, Tanser F, Vanneste A . Hlabisa DSS,South Africa: Population, Health and Survival at INDEPTH Sites. Ottawa:IDRC, 2002: p. 213–20.
51. Tanser, F., Hosegood, et al. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. Int. J. Epidemiology, 2007: p. 1-6.
52. Botha, J.L., & Bradshaw, D. African vital statistics-a black hole? South African Medical Journal 1985. 67 p. 977-981.
53. Yach, D. & Buthelezi, G. Health Status, in South African Health Review. 1995, Health Systems Trust and the Henry J.Kaiser Family Foundation,Durban.
54. Kahn K.; Tollman S.M.; Garenne, M.; Gear, J.S. Validation and application of verbal autopsies in a rural area of South Africa. Trop Med Int Health, 2000. 5(11): p. 824-31.

55. Muhwava,W. & Nyirenda, M. Demographic and Socio-Economic Trends in the ACDIS. 2008, Monograph Series No. 2. Africa Centre for Health and Population Studies, Somkhele,South Africa
56. Boscoe, F., McLaughlin, C., Schymura MJ, Kielb L. Visualization of the spatial scan statistic using nested circles.Health & Place, 2003.9(3):273–277.
57. Hosegood, V., Vanneste, A-M.,Timaeus I.M . Levels and causes of adult mortality in rural South Africa. AIDS, 2004. 18: p. 663-671.
58. Herbst, K., Cooke, G., Barnighausen T, Kany-Kany A, Newel M-L. Early impact of on adult mortality from a government ART programme in rural KwaZulu -Natal, 2008.
59. Wakefield, J., Kelsall, J.E., Morris, S.E. Clustering, cluster detection. Clustering, cluster detection, and spatial variation in risk. : In: Elliott P, Wakefield JC, Best NG, et al, eds. Spatial epidemiology: methods and applications. New York, NY: Oxford University Press, 2002:128–52.
60. Kulldorff, M., Athas, WF., Feurer EJ, et al. , Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. . Am J Public Health, 1998. 88: p. 1377-80.
61. Kulldorff, M., Feuer, E.J., Miller, B.A. Freedman, L.S, Breast cancer clusters in the northeast United States: a geographic analysis. Am J Epidemiol, 1997. 146: p. 161-170.
62. Michaud, J., Ellen, J.M., Johnson, S.M, et al. Responding to a community outbreak of syphilis by targeting sex partner meeting location: an example of a risk-space intervention. Transm Dis 2003. 30: p. 533-8.
63. Besag, J. & Newell J . *The detection of clusters in rare diseases.* J Roy Stat Soc,Ser A, 1991. 154(143-155).
64. Lawson, A.B. Statistical Methods in Spatial Epidemiology, in Wiley Series in Probability and Statistics. 2001, John Wiley & Sons. pp.298.
65. Morrison, A., Getis, A., Santiago, M, Rigau-Perez J.G, Reiter, P. Exploratory space-time analysis of reported dengue cases during an outbreak in Florida, Puerto Rico, 1991-1992. Am J Trop Med Hyg 1998. 58: p. 287-298.
66. Getis, A., Morrison, A.C, Gray K, Scott, T.W. Characteristics of the spatial pattern of the dengue vector, Aedes aegypti, in Iquitos, Peru. Am. J. Trop. Med. Hyg, 2003. 6: p. 494-505.
67. Kitron, U., Michael, J., Swanson J., Haramis, L. Spatial analysis of the distribution of LaCrosse encephalitis in Illinois, using a geographic information system and local and global spatial statistics. Am J Trop Med Hyg, 1997. 57(469-75).
68. Chadee, D.D., Kitron U. Malaria surveillance in Trinidad: imported cases and risk of outbreaks. Am J Trop Med Hyg, 1999. 61: p. 513-517.
69. Hosegood V, Vanneste AM, Timaeus IM. Levels and causes of adult mortality in rural South Africa: the impact of AIDS. AIDS 2004;18:663–71.
70. Jacky, M.J., Frank, C.C., David,C. & Jonathan, M.E. Geographic identification of high transmission areas in Baltimore.Am J Epidemiol 2005;161:73-80
71. Agricola, O., Wayne, M.S., Pascal, M., Dean, M., John, H. & Jeff, W. Investigation of clusters of giardiasis using GIS and Spatial scan statistic.IJHG 2004,3:1-11

72. Tanser,F.C "The application of GIS technology to equitably distribute fieldworker workload in a large, rural South African health survey," Tropical Medicine & International Health Vol. 7, No.1 (2002), pp. 80–90.
73. McLafferty, "GIS and health care."
74. Janet,M,Wojcicki. Socio-Economic Status as a risk factor for HIV infection in women in East,Central and Southern Africa: A Systemmatic Review. *J. biosoc. Sci.* (2005), **37**, 1–36_ 2004 Cambridge University Press
75. Raphael, I. Trends in HIV Prevalence and Incidence in Magu DSS, 1994-2007. *In program and Abstracts: The 8th INDEPTH Network Annual General & Scientific Meeting, p46. Dar es Salaam, Tanzania, 22-26th September 2008.*
76. Mark, U. Estimating need for Anti-Retroviral Treatment: a new method developed using data from Magu DSS. *In program and Abstract*s: *The 8th INDEPTH Network Annual General & Scientific Meeting, p45. Dar es Salaam, Tanzania, 22-26th September 2008.*
77. Knight, R. Population and HIV/AIDS, Statistics South Africa, 2006. www.statssa.gov.za accessed August 20, 2009
78. World Health Organization. Mortality Database 2003. http://www.who.int/whosis/. Accessed: 2009-08-20.
79. Timæus, I. M. Impact of the HIV epidemic on mortality in sub-Saharan Africa: evidence from national surveys and censuses, 1998. AIDS, 12(supplement 1):S15-S27.

UNIVERSITY OF
KWAZULU-NATAL

Research Office
BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Nelson R Mandela School of Medicine
Private Bag 7, Congella 4013
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769
Fax: 27 31 2604609
Email: buccas@ukzn.ac.za
Website: www.ukzn.ac.za

21 December 2006

**Professor A J Herbst**
Africa Centre

**Fax : 035 550 7565**

**PROTOCOL: A socio-demographic platform for population-based reproductive health research in a rural health district of KwaZulu-Natal. (GC Solarsh,** Paediatrics) A J Herbst, Africa Centre. Ref: E009/00

## RECERTIFICATION RATIFICATION NOTICE

Dear Professor A J Herbst

Further to our letter to you dated 29 November 2006, this letter serves to notify you that at a full sitting of the Biomedical Research Ethics Committee Meeting held on **12 December 2006**, the Committee RATIFIED the sub-committee's decision to approve the Recertification of the above study dated 13 November 2006.

Yours sincerely

SURAIYA BUCCAS
Ethics Research Administrator

SB/sn