

https://dspace.org/

The estimation of AIDS-related deaths amongst the black population of South Africa and the impact on the ASSA AIDS model

Mitchell, Julie-Anne

2014-05-30
http://hdl.handle.net10539/14724

Downloaded from a DSpace repository. https://dspace.org/

The Estimation of AIDS-Related Deaths Amongst the Black Population of South Africa and the Impact on the ASSA AIDS Model

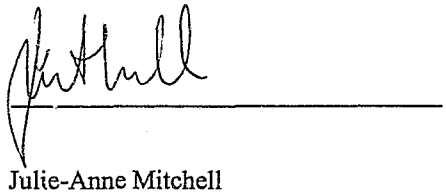
Julie-Anne Mitchell

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science

Johannesburg, 1998

DECLARATION

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.



Julie-Anne Mitchell
March 1998

ABSTRACT

This dissertation reviews the AIDS epidemic in South Africa. The mathematical structures of various AIDS models are compared. The model produced by the Actuarial Society of South Africa is outlined in detail as it is accessible, of recent origin and in use. Existing models are based on HIV prevalence data only. Projected deaths may therefore be inaccurate. AIDS-related deaths occurring within the black population are estimated from CSS reports, with adjustments for underreporting. Between 10 000 and 15 000 and between 20 000 and 25 000 AIDS-related deaths appear to have occurred during 1993 and 1994 respectively. Analysis by recorded cause of death and occupation group is performed to formulate the profile an AIDS-related death. KwaZulu-Natal deaths were estimated separately. Estimates are tentative due to poor data quality. The ASSA model was calibrated against the estimates by changing specific parameters, and projections were revised. Finally, the calibrated model was applied to a typical insurance problem.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Anthony Asher, for his fresh perspectives and unfailing support. The priority given to this dissertation and time devoted to it are much appreciated.

Many thanks to Greg Solomon for his keen interest in the subject, for keeping me informed of recent developments, and for providing particular references.

I am grateful to Professor Rob Dorrington for his invaluable opinion on the reliability of South African data and the necessary adjustments for underreporting. Thanks also to Alan Whitelock-Jones for detail on the ASSA model.

Appreciation is extended to Sulaiman Bah and Alida Casteleijn of the CSS for providing South African death and population data. Thanks also to John Brombe[illegible] for insurance data.

The medical perspective provided by Dr Freddy Sitas is much appreciated.

Finally, on a more personal note, I would like to thank Robert for keeping my chin up and keeping me focused on the ultimate goal.

CONTENTS

DECLARATION ii
ABSTRACT iii
ACKNOWLEDGEMENTS iv
LIST OF FIGURES x
LIST OF TABLES xv

1. Introduction 1

1.1 Background 1
1.2 Statement of the Problem 2
1.3 Dissertation Objectives 2

2. The Epidemic in South Africa 4

2.1 AIDS: The Disease 4
2.1.1 History 4
2.1.2 The HI-Virus 5
2.1.3 Progression of the Disease 6
2.1.4 Transmission of the HI-Virus 7
2.1.5 Patterns of Spread 9
2.1.6 Impact 10
2.2 The South African Position 11
2.2.1 The South African Disease 11
2.2.2 The Current Situation 12

3. An Overview of AIDS Models in South Africa 17

...und to Mathematical Modelling 17
...ucture of AIDS Models 17
Criterion 1: Level of Detail 18
Criterion 2: Mathematical Structure 18

3.2.3 Criterion 3: Calibration 19
3.3 The Model Produced by Edelston 19
3.3.1 Doomsday Scenario 20
3.3.2 Super-Optimist Scenario 21
3.4 The Model Produced by Padayachee and Schall 21
3.4.1 Extrapolation 21
3.4.2 The Direct Method 22
3.4.3 Back Calculation 22
3.5 The Model Produced by Schall 23
3.5.1 Scenario 1 24
3.5.2 Scenario 2 24
3.5.3 Scenario 3 24
3.6 The Model Produced by Doyle 25
3.7 The Model Produced by Groeneveld and Padayachee 26
3.8 The Model Produced by Gregson *et al* 28
3.9 The Model Produced by Bos and Bulatao 28
3.10 The Model Produced by Chin and Lwanga 29
3.11 The Model Produced by Bongaarts 30
3.12 Other Models Produced Overseas 31
3.13 Discussion and Comparison 32
3.14 Summary 33

4. The ASSA Model in Detail 35

4.1 Background 35
4.2 An Overview of the Structure of the Model 36
4.3 Inputs 37
4.3.1 Notation 37
4.3.2 Starting Population 39
4.3.3 HIV-Infected Immigrants 39
4.3.4 Fertility 39
4.3.5 Non-AIDS Mortality Rates 40
4.3.6 AIDS Mortality Rates 40
4.3.7 Infection 41
4.4 Projections 43
4.4.1 Probability of Infection 44
4.4.2 HIV-Negative Individuals 45
4.4.3 HIV-Positive Individuals 46
4.4.4 Non-AIDS Deaths 47
4.4.5 AIDS-Related Deaths 47
4.5 Summarised Results 48
4.6 Comparison with the Doyle Model 53
4.7 Summary 54

5. Principal Estimation of AIDS-Related Deaths 55

5.1 Overview of Method 55
5.2 Notation 56
5.3 Estimation of the Actual Number of Black Deaths Which Occurred in 1991, 1992, 1993 and 1994 58
5.3.1 Extraction of Black Reported Deaths from Total Reported Deaths 59
5.3.2 The Preston and Coale Method 67
5.3.3 The Bennett and Horiuchi Method 71
5.3.4 Application of the Preston and Coale and the Bennett and Horiuchi Methods 72
5.3.5 Comparison and Discussion of Results from the Preston and Coale and the Bennett and Horiuchi Methods 78
5.3.6 The Estimated Number of Actual Black Deaths 79
5.4 Estimation of the Size and Demographic Structure of the Black Population 80
5.4.1 Available Data 80
5.4.2 Comparison of the 1991 and Preliminary 1996 Census Results 81
5.4.3 Method Used to Estimate the Black Population 86
5.4.4 Results 88
5.4.5 Possible Distortions 89
5.5 Estimation of Actual Mortality Rates for the Black Population 90
5.6 Selection and Adjustment of Expected Mortality Rates for the Black Population 93
5.6.1 Selection 93
5.6.2 Adjustment 96
5.6.3 Problems 99
5.7 The Estimated Number of Black AIDS-Related Deaths 99
5.7.1 Discussion 102
5.8 Summary 104

6. Further Analysis of AIDS-Related Deaths 106

6.1 Analysis of AIDS-Related Deaths by Recorded Cause of Death 106
6.1.1 Adjustment for Underreporting 110
6.1.2 Analysis of Cause-Specific Mortality Rates 112
6.1.3 Results 118
6.1.4 Discussion 119
6.2 Analysis of AIDS-Related Deaths by Occupation 120
6.2.1 Adjustment for Underreporting 123
6.2.2 Estimation of the Population by Occupation Group 123
6.2.3 Analysis of Mortality Rates by Occupation 124
6.2.4 Results 131

6.2.5 Discussion 132
6.3 Analysis of AIDS-Related Deaths in KwaZulu-Natal 133
6.3.1 Adjustment for Underreporting 134
6.3.2 Estimation of the Population of KwaZulu-Natal 134
6.3.3 Analysis of Mortality Rates in KwaZulu-Natal 135
6.3.4 Results 137
6.3.5 Discussion 138
6.4 Summary 138

7. Calibration of the ASSA Model 140

7.1 Adjustment of Demographic Assumptions Underlying the ASSA Model 140
7.1.1 Adjustment of Age and Gender Distribution 140
7.1.2 Adjustment of Total Fertility Rates 142
7.1.3 Adjustment of Population Size 143
7.2 Calibration Against Estimated AIDS-Related Deaths 145
7.2.1 Comparison of the Adjusted ASSA Model and Empirical Data 145
7.2.2 Calibration Method 146
7.2.3 The Calibrated Model 148
7.3 Goodness-of-Fit Tests 150
7.3.1 Test 1 151
7.3.2 Test 2 151
7.3.3 Conclusion 152
7.4 Revised Projections 155
7.5 Discussion 155
7.5.1 Comparison with the Original ASSA Model 155
7.5.2 Comparison with Empirical Data 156
7.6 Summary 157

8. Application to the Insurance Industry 159

8.1 Preliminaries 159
8.2 Policies Written Prior to HIV-Testing 160
8.3 Policies Written Subsequent to HIV-Testing 160
8.4 Example 161
8.4.1 Statement of the Problem 161
8.4.2 Solution 161
8.5 Discussion 162
8.6 Summary 162

9. Conclusion 164

APPENDICES 166

REFERENCES 174
BIBLIOGRAPHY 185

LIST OF FIGURES

Figure 2.1 The age and gender distribution of reported AIDS cases as at 30 November 1995 15

Figure 2.2 The age and gender distribution of AIDS-related deaths reported during 1994 15

Figure 4.1 Sexual activity indices assumed in th SSA model 43

Figure 4.2 Comparison of the antenatal clinic data published by the Department of Health, and the HIV prevalence of pregnant women as projected by the ASSA model 48

Figure 4.3 Projected normal and AIDS-related deaths 50

Figure 4.4 Projected female mortality curves, with and without the effect of AIDS 51

Figure 4.5 Projected male mortality curves, with and without the effect of AIDS 51

Figure 4.6 Projected HIV prevalence rates amongst females aged 15-45 in the PRO, STD and RSK risk groups 52

Figure 4.7 Projected HIV prevalence rates amongst males aged 15-45 in the PRO, STD and RSK risk groups 52

Figure 5.1 Extrapolated standard male mortality curves 60

Figure 5.2 Extrapolated standard female mortality curves 61

Figure 5.3 Adjusted standard male mortality curves 64

Figure 5.4 Adjusted standard female mortality curves 65

Figure 5.5 Scenario 1: Plot of $c_{t,4,s,x^3}$ against x for the black population using the Preston and Coale method 74

Figure 5.6 Scenario 1: Plot of $c_{t,4,s,x^5}$ against x for the black population using the Bennett and Horiuchi method 75

Figure 5.7 Scenario 2: Plot of $c_{t,4,s,x^3}$ against x for the black population using the Preston and Coale method 75

Figure 5.8 Scenario 2: Plot of $c_{t,4,s,x^5}$ against x for the black population using the Bennett and Horiuchi method 76

Figure 5.9 Scenario 3: Plot of $c_{t,4,s,x^3}$ against x for the black population using the Preston and Coale method 76

Figure 5.10 Scenario 3: Plot of $c_{t,4,s,x^5}$ against x for the black population using the Bennett and Horiuchi method 77

Figure 5.11 Total fertility rates for black women used in Sadie's model 82

Figure 5.12 Udjo's estimated total fertility rates 82

Figure 5.13 Estimated actual mortality rates for black males under Scenario 2 90

Figure 5.14 Estimated actual mortality rates for black females under Scenario 2 91

Figure 5.15 Estimated actual mortality rates for black males under Scenario 3 91

Figure 5.16 Estimated actual mortality rates for black females under Scenario 3 92

Figure 5.17 Estimated actual mortality curves for the black population in 1994 92

Figure 5.18 Comparison of various male African mortality rates and the 1991 actual rates under Scenario 2 95

Figure 5.19 Expected mortality rates for the black population 96

Figure 5.20 Estimated number of black AIDS-related deaths under Scenarios 2 and 3 101

Figure 5.21 The age structure of correctly reported AIDS-related deaths in 1992, 1993 and 1994 combined 103

Figure 6.1 Estimated completeness of reporting of deaths in all population groups 112

Figure 6.2 Estimated actual rates of mortality from tuberculosis (TB) 113

Figure 6.3 Estimated actual rates of mortality from other infectious and parasitic diseases (INF), for males 114

Figure 6.4 Estimated actual rates of mortality from other infectious and parasitic diseases (INF), for females 114

Figure 6.5 Estimated actual rates of mortality from neoplasms (NEO), for males 115

Figure 6.6 Estimated actual rates of mortality from neoplasms (NEO), for females 115

Figure 6.7 Estimated actual rates of mortality from endocrine, nutritional and metabolic diseases and immunity disorders (END), for males 116

Figure 6.8 Estimated actual rates of mortality from endocrine, nutritional and metabolic diseases and immunity disorders (END), for females 116

Figure 6.9 Estimated actual rates of mortality from symptoms, signs and ill-defined conditions (ILL) 117

Figure 6.10 Estimated actual rates of mortality for: professional, semi-professional and technical occupations 125

Figure 6.11 Estimated actual rates of mortality for: management, executive and administrative occupations 126

Figure 6.12 Estimated actual rates of mortality for: clerical and sales occupations 126

Figure 6.13 Estimated actual rates of mortality for: transport, delivery and communications occupations 127

Figure 6.14 Estimated actual rates of mortality for: service occupations 127

Figure 6.15 Estimated actual rates of mortality for: farming and

related occupations 128

Figure 6.16 Estimated actual rates of mortality for: artisan, apprentice and related occupations 128

Figure 6.17 Estimated actual rates of mortality for: production supervisor, miner, quarry and related worker 129

Figure 6.18 Estimated actual rates of mortality for: occupation unspecified and not elsewhere classified 129

Figure 6.19 Estimated actual rates of mortality for: total not economically active 130

Figure 6.20 Estimated actual rates of mortality amongst males living in KwaZulu-Natal 136

Figure 6.21 Estimated actual rates of mortality amongst females living in KwaZulu-Natal 136

Figure 7.1 Population pyramid for the starting population in the unadjusted ASSA model 144

Figure 7.2 Population pyramid for the starting population in the adjusted ASSA model 144

Figure 7.3 Curves of the proportion of HIV-infected individuals surviving by duration in the original and calibrated ASSA models 148

Figure 7.4 Sexual activity indices assumed in the original and calibrated ASSA models 149

Figure 7.5 Comparison of the antenatal clinic data published by the Department of Health, and the HIV prevalence of pregnant women as projected by the calibrated model 150

Figure 7.6 Revised projected HIV prevalence rates amongst females aged 15-45 in the PRO, STD and RSK risk groups 152

Figure 7.7 Revised projected HIV prevalence rates amongst males aged 15-45 in the PRO, STD and RSK risk groups 153

Figure 7.8 Revised projected normal and AIDS-related deaths 153

Figure 7.9 Revised projected female mortality curves, with and

without the effect of AIDS 154

Figure 7.10 Revised projected male mortality curves, with and without the effect of AIDS 154

LIST OF TABLES

Table 2.1 Observed national HIV prevalence rates amongst women attending antenatal clinics in South Africa 13

Table 2.2 Observed provincial HIV prevalence rates amongst women attending antenatal clinics in South Africa 14

Table 2.3 Observed HIV prevalence rates amongst women attending antenatal clinics in South Africa, by age 14

Table 3.1 The number of HIV-infected individuals in South Africa, as projected by Edelston 20

Table 3.2 HIV prevalence rates for various subgroups of the black population aged 15-49, as projected by Padayachee and Schall 22

Table 3.3 The number of HIV-infected blacks aged 15-49, as projected by Padayachee and Schall 22

Table 3.4 The number of HIV-infected individuals and AIDS-related deaths, as projected by Doyle 26

Table 3.5 The number of newly HIV-infected individuals and cumulative AIDS cases amongst black adults, as projected by Groeneveld and Padayachee 27

Table 4.1 Values of Ω_{r_i,r_j} assumed in the ASSA model 41

Table 4.2 Projected statistics indicating the course of the AIDS epidemic in South Africa 49

Table 4.3 Projected HIV prevalence rates amongst various subgroups of the population aged 15-45 50

Table 5.1 Deaths grouped by year of death and year of registration 58

Table 5.2 Estimated number of reported black deaths 66

Table 5.3 Estimated completeness of reporting of black deaths using the Preston and Coale (P & C) and the Bennett and Horiuchi (B & H) methods 77

Table 5.4 Estimated number of actual black deaths 80

Table 5.5 Estimated size of the black population aged 0 to 75 89

Table 5.6 Estimated values of $R_{t,4,S}$ used to adjust the expected mortality rates 98

Table 5.7 Estimated number of black AIDS-related deaths under Scenario 2 101

Table 5.8 Estimated number of black AIDS-related deaths under Scenario 3 102

Table 6.1 The number of recorded deaths amongst individuals aged 20 to 55 with AIDS as the underlying cause of death 108

Table 6.2 Breakdown of estimated adult, male AIDS-related deaths by recorded cause of death 118

Table 6.3 Breakdown of estimated adult, female AIDS-related deaths by recorded cause of death 119

Table 6.4 "Excess" adult deaths and percentage increase in deaths since 1991 by occupation group 131

Table 6.5 Estimated number of AIDS-related deaths in KwaZulu-Natal 137

Table 7.1 Comparison of the projected and estimated actual population aged 0 to 75 (in thousands), assuming a starting population of 25,2 million 145

Table 7.2 Comparison of the estimated actual AIDS-related deaths occurring in 1993 and 1994 and those projected by the adjusted ASSA model 145

Table 7.3 Comparison of the estimated actual AIDS-related deaths occurring in 1993 and 1994 and those projected by the

calibrated ASSA model 149

Table 7.4 Revised projected statistics indicating the course of the AIDS epidemic in South Africa 155

Table A1 Estimated number of reported black deaths under Scenario 2 167

Table B1 Estimated number of actual black deaths under Scenario 2 168

Table C1 Estimated black population aged 0 to 75 under Scenario 2 169

Table D1 Age and gender distributions in the unadjusted ASSA model 170

Table D2 Age and gender distributions in the adjusted ASSA model 171

Table E1 Mortality rates projected by the calibrated ASSA model - 2000 172

Table E2 Mortality rates projected by the calibrated ASSA model - 2005 173

1

Introduction

1.1 Background

While the AIDS epidemic has caused concern throughout the world, the most rapid spread is occurring in developing countries. Approximately two-thirds of infected individuals are estimated to be living in Sub-Saharan Africa, where heterosexual transmission is the dominant pattern of spread (UNAIDS, 1997). Consequences include a reduction in life expectancy, rapid increase in the number of orphans, disruption of traditional responsibilities within the family, increasing demands on health care, productivity losses, and reduced economic growth (Broomberg *et al*, 1991; Cross, 1993; Whiteside, 1993; UN, 1994; and Patel, 1995).

The South African epidemic is at an earlier stage than the rest of Africa. A recent HIV prevalence survey indicated that 14% of antenatal clinic attendees were HIV-positive in 1996. By extrapolating to the rest of the population, it is estimated that approximately 6% of the South African population and 11% of adults were HIV-infected during 1996 (*Business Day*, 1997 and *The Star*, 1997).

A more comprehensive discussion of the AIDS epidemic and its current state in South Africa is provided in Chapter 2.

A number of models have been developed to project the spread of the epidemic in South Africa (Edelston, 1988; Doyle and Millar, 1990; Padayachee and Schall, 1990; Schall, 1990; Groeneveld and Padayachee, 1992; ASSA, 1996; and Gregson *et al*, 1997). Other models were developed overseas to project the epidemic in different regions throughout the world (Bos and Bulatao, 1992; Chin and Lwanga, 1991; and Bongaarts, 1989). Chapter 3 outlines and contrasts the different projection methods, and results are reconciled where possible. The projections provide information for decisions on budget allocation, health care provision, and AIDS-awareness campaigns, amongst others.

The increase in morbidity and mortality caused by the AIDS epidemic is a source

of concern to the life insurance industry, which requires mortality projections for product pricing and reserving. As a result, the actuarial profession has developed a number of models to project future AIDS-related deaths.

The Actuarial Society of South Africa recently published a model for the recommendation of mortality guidelines for the actuarial profession (ASSA, 1996). Its parameters may be adjusted, and the model can thus be used to predict the impact on a particular life office. This model is discussed in detail, the mathematical structure is clarified, and summarised results are shown in Chapter 4.

1.2 Statement of the Problem

The existing models are generally calibrated against HIV prevalence data, rather than data on AIDS cases or AIDS-related deaths. HIV prevalence data are readily available from annual antenatal clinic surveys, and present a picture of the current state of the epidemic based on a sample of pregnant women. Reliable data on AIDS-related deaths are particularly difficult to obtain, for two reasons: (i) a substantial proportion of deaths are never reported (Botha and Bradshaw, 1985; Dorrington, 1989; Dorrington, Martens and Slawski, 1991; Bah and Kleinschmidt, 1997); and (ii) AIDS-related deaths which are reported may be incorrectly recorded with respect to cause of death.

Since models have not been calibrated against data on AIDS-related deaths, projected deaths may be inaccurate. The projection of deaths is critically important to actuaries, and imprecise and untested models are a major cause for concern.

1.3 Dissertation Objectives

This dissertation aims to address the problem by estimating AIDS-related deaths. The Central Statistical Services (CSS) publishes an annual report on recorded deaths. These reports are likely to provide some clue as to the impact of AIDS on population mortality, although require adjustment to overcome the data inadequacies. Appropriate techniques are used to adjust for underreporting (Preston *et al*, 1980; Bennett and Horiuchi, 1981).

Mortality rates have historically followed a relatively well-defined and stable pattern. The AIDS epidemic is expected to alter the shape of mortality curves. Actual mortality rates are estimated and compared against suitable expected rates, and changes in the shape of the mortality curve attributed to AIDS.

The estimation of AIDS-related deaths is tackled in Chapter 5. Recorded deaths

are adjusted for underreporting using two alteration approaches. The estimates are analysed further in Chapter 6, with respect to recorded cause of death and occupation group of the deceased. The number of AIDS-related deaths occurring in KwaZulu-Natal is also estimated to provide some indication of the outlook for areas with a less advanced epidemic.

The dissertation aims to calibrate the ASSA model against the estimated AIDS-related deaths. This is performed by changing relevant parameters as described in Chapter 7. Statistical methods for testing goodness of fit are applied, and highlight the uncertainty surrounding the estimates. Revised model projections are given, and may be compared against the original projections and empirical data. Such calibration should result in a model which is more reflective of reality. Decision-makers are likely to be more confident in recommendations which are based on such a model.

Chapter 8 applies the calibrated model to the estimation of mortality rates for groups of assured lives. The importance of accurate models of the AIDS epidemic for the insurance industry is thus emphasised.

2

The Epidemic in South Africa

This chapter is divided into two major sections: section 2.1 which discusses the disease itself, and section 2.2 which places the epidemic in a South African context. The literature on the disease is extensive, and includes Barnett and Blaikie, 1992; Cross and Whiteside, 1993; Schoub, 1994; WHO, 1994; and Lachman, 1995, as listed in the Bibliography.

2.1 AIDS: The Disease

2.1.1 History

In 1980, US doctors became aware of an increase in the incidence of an unusual form of pneumonia (*Pneumocystis carinii*, or PCP), as well as an uncommon tumour called Kaposi's sarcoma (KS), amongst homosexual men. PCP had previously been found in patients with severely suppressed immune systems, and KS had become increasingly evident amongst patients on immunosuppressive drug therapy. Thus, the early indications pointed to a disease causing severe immunosuppression amongst homosexual men - hence the original name "GRIDS" (Gay-Related Immune Deficiency Syndrome).

Characteristics of this new disease began appearing amongst haemophiliacs and patients who had received blood transfusions. Intravenous drug users also began showing symptoms of the disease. The evidence now suggested that the infectious agent was carried in the blood, and was not restricted to homosexuals. The United States Centres for Disease Control officially coined the name "AIDS" (Acquired Immunodeficiency Syndrome) in 1982. In 1983, Belgian and French doctors diagnosed AIDS amongst a number of their male and female African patients. These patients were heterosexual and did not use drugs, and heterosexual transmission was suspected.

The infectious agent was later identified as a virus, termed the human immunodeficiency virus (abbreviated HIV). A variant of the first HIV was later discovered, and named HIV-2. The majority of AIDS cases resulting from HIV-2 have arisen in West Africa.

The World Health Organisation (WHO) established a Global Programme on AIDS in order to monitor the epidemic, set global priorities, and obtain international co-operation for the control of the disease. The number of AIDS cases reported to the WHO increased dramatically each year. By 1983, just under 5 000 cases had been reported. By 1986, over 45 000 cases had been reported (WHO, 1988). Extreme levels of under-reporting were suspected, particularly in the third world countries.

The latest estimates published by the Joint United Nations and World Health Organisation Programme on HIV/AIDS (UNAIDS, 1997) indicate that during 1997, approximately 5,8 million people became HIV-infected, 30,6 million people were living with HIV/AIDS, and 2,3 million AIDS-related deaths occurred. 11,7 million AIDS-related deaths are estimated to have occurred since the start of the epidemic.

2.1.2 The HI-Virus

The HI-virus is a member of the family of retroviruses. Retroviruses invade host cells, and insert a copy of their own DNA into the host cells' DNA. The hosts' reproductive resources are then redirected to the production of further HI-viruses.

The HI-virus invades cells known as T-cells (also known as T- lymphocytes or CD4+ cells). T-cells play an important role in the immune system by stimulating B-cells to produce antibodies. These antibodies eliminate foreign particles and organisms within the body, thus fighting illness. By invading and subsequently destroying the T-cells, the HI-virus gradually attacks the immune system. This opens the door to a number of opportunistic infections, which otherwise pose little threat to a healthy immune system. HIV-infected individuals, with their compromised immune systems, usually die from these opportunistic infections.

The HI-virus also attacks a number of other cells directly, particularly those involved in the central nervous system, gut, blood forming elements, kidneys, joints and skin.

HIV-2 differs slightly from the more common HIV-1 virus with respect to genetic composition. Research suggests that individuals infected with HIV-2 progress more slowly towards AIDS.

2.1.3 Progression of the Disease

Clinicians differ in opinion as tc the number of distinct stages which can be identified in the progression of HIV-infection. Five stages will be identified in this dissertation.

Stage 1: Window Period

An individual in the window period is HIV-positive, but has no HIV antibodies. Laboratory tests which test for these antibodies would thus be unable to detect the presence of HIV. The individual is, however, highly infectious during this stage. Fever-like symptoms may become manifest, although this is uncommon amongst children. The duration of the window period is uncertain, but may be around six to eight weeks (Schoub, 1994).

Stage 2: Latent Phase

This stage begins when HIV antibodies are produced (i.e. the individual becomes "seropositive"). An individual in the latent phase is infectious, although is usually unaware of the infection since no illness is experienced. During this time, the HI-virus is replicating and beginning its attack the immune system. The latent phase can last for 8 to 15 years (Schoub, 1994).

Stage 3: Progressive Generalised Lymphadenopathy (PGL)

The lymph tissue is usually the first tissue to show signs of HIV infection. During this phase, the patient becomes aware of swollen lymph glands (lymphadenopathy), especially in the head and neck region. The symptoms generally persistent for more than three months, and are an important clinical indication of HIV infection.

Stage 4: AIDS-Related Complex (ARC)

Individuals in this phase show a number of symptoms which precede the advent of full-blown AIDS, such as weight loss, diarrhoea, and recurrent thrush. The immune system shows signs of severe immunosuppression.

Stage 5: AIDS

In this stage, opportunistic infections occur with increasing severity and frequency. Common opportunistic infections occurring during this stage include

tuberculosis, severe weight loss coupled with diarrhoea (known as "slim disease" in Africa), PCP, *herpes zoster*, AIDS dementia, and KS. The individual is highly infectious during this stage. The direct and indirect (via damage to the immune system) consequences of HIV damage become increasingly severe, leading to death within 1,5 to 2 years, on average (Schoub, 1994).

Not all HIV-infected individuals pass through all five stages mentioned above. In particular, patients may progress from the latent phase to the ARC phase directly, without evidence of lymphadenopathy. It is also uncertain whether all HIV-infected individuals will progress to full-blown AIDS. However, if an individual does so, death from one of the infections is inevitable.

The time from initial HIV infection to the manifestation of AIDS is called the incubation period. The average incubation period for an HIV-1-infected adult is believed to be between 8 and 13 years, and between 10 and 15 years for an HIV-2-infected adult (Schoub, 1994). The incubation period appears to be shorter for children. Evidence has emerged suggesting that older individuals progress faster and experience higher mortality rates (Phillips *et al*, 1994).

Other sets of clinical criteria have been established in order to produce an objective definition of AIDS. These classification systems generally involve HIV testing, the demonstration of certain "indicator diseases" which are indicative of immunosuppression, and specified manifestations of AIDS (including direct consequences of HIV damage).

The number of CD4+ cells per microlitre of blood gives some indication of the state of an individual's immune system. People with healthy immune systems usually have a CD4+ cell count of approximately 950. This drops during the window period, recovers slightly during the latent phase, and then declines as the HIV infection progresses. People with AIDS generally have a CD4+ count of less than 200. This value is used in the USA in the definition of AIDS.

The concentration of the virus in the body can also be measured over the course of the infection. High concentrations are found during the window period. The concentration then drops during the latent phase, and then rises as the infection progresses. An infected individual is thus most infectious at the very early and terminal stages of infection.

2.1.4 Transmission of the HI-Virus

The HI-virus has been isolated from most body fluids, such as blood, semen, vaginal secretions, breast milk, cerebrospinal fluid, tears, and saliva. However, it can only be regularly isolated from the first three. The HI-virus is thus transmitted via an exchange of these body fluids, and in a manner which limits its exposure to a hostile environment. Transmission may occur through single exposure to a high

concentration of the HI-virus, or through recurrent and repetitive exposures to low concentrations. The main routes of transmission are described below.

Sexual Transmission

Unprotected sexual intercourse with an infected partner exposes the uninfected partner to the risk of HIV infection.

Sexual intercourse between heterosexuals is the most common method of transmission, although the risk per contact is fairly low at 0,1% to 1,0% per exposure (*AIDS Analysis Africa (Southern Africa Edition)*, 1995). In the absence of other co-factors, females are more vulnerable to infection from an HIV-infected male, than vice versa. Relative risks of infection of between 2:1, and 10:1 have been quoted (Schoub, 1994; *AIDS Analysis Africa (Southern Africa Edition)*, 1995 and Solomon, 1995). Females in their teens and those who have reached menopause appear to be more susceptible to infection.

Co-factors which tend to increase the risk of HIV infection include the presence of sexually transmitted diseases (notably those causing genital ulcers), the absence of circumcision in males, as well as a suppressed immune system (due to drugs, malaria, tuberculosis, oral contraceptives, etc.). As discussed in section 2.1.3, the risk of infection varies with the stage of infection. Sexual intercourse with an infected individual in the window period or at the terminal stages of infection thus poses a greater risk.

Rectal intercourse between male homosexuals is believed to expose the uninfected partner to a probability of infection of between 0,1% and 4% (*AIDS Analysis Africa (Southern Africa Edition)*, 1995). Sexual contact between female homosexuals plays no role in the transmission of HIV.

Vertical Transmission

Vertical transmission is the name given to the transmission from an infected mother to her infant. This can occur at three stages: whilst the foetus is in the womb (intrauterine infection), during birth (perinatal infection), or during breast-feeding (being less common). The overall risk of vertical transmission is believed to be 30% (*AIDS Analysis Africa (Southern Africa Edition)*, 1995 and FitzSimons, 1993), but may depend on the stage of infection of the mother. When the mother has AIDS, the risk of infection increases to about 80% (Schoub, 1994).

Transmission by Blood Transfusion and Exchange of Blood Products

This route of transmission poses a risk of infection of more than 90% (*AIDS*

Analysis Africa (Southern Africa Edition), 1995). However, the screening of blood has greatly reduced the incidence of transfusion-acquired HIV in the developed world as well a large number of developing countries.

Transmission by Shared Needles and Syringes

Intravenous drug users often share hypodermic needles or syringes used to habitually inject drugs into their veins. An uninfected individual may become infected when sharing needles or syringes with an infected individual. The probability of such infection is believed to be between 0,5% and 1% (*AIDS Analysis Africa (Southern Africa Edition)*, 1995).

Intravenous drug users often suffer from infections which suppress the immune system (e.g. hepatitis B, cytomegalovirus, etc.), as well as sexually transmitted diseases. Susceptibility to HIV infection is thus increased.

Transmission by Injury

Members of the medical profession may be exposed to infected blood through accidental injury caused by needles and sharp instruments. The probability of transmission via this route is surmised to be less than 0,5% (*AIDS Analysis Africa (Southern Africa Edition)*, 1995). However, coupled with the infrequency of such injuries, this route of transmission is relatively insignificant. Only 41 cases of infection by needle-stick injury had been reported world-wide by 1994 (Schoub, 1994).

2.1.5 Patterns of Spread

Three distinct patterns of spread have been identified (Chin and Mann, 1988):

Pattern I

This pattern is prevalent in the Western world, where most cases are amongst homosexual or bisexual men and intravenous drug users. Other routes of transmission are uncommon, and the HIV prevalence within the general population is low. For example, the Joint United Nations and World Health Organisation Programme on HIV/AIDS estimated that 0,3% and 0,6% of adults living in Western Europe and North America respectively were HIV-positive as at December 1997 (UNAIDS, 1997).

Pattern II

This pattern is prevalent in Sub-Saharan Africa and parts of Asia. Heterosexual intercourse is the most common route of transmission, followed by vertical transmission and transmission via contaminated blood. HIV prevalence within the general population is relatively high. For example, 7,4% of adults living in Sub-Saharan Africa were estimated to be HIV-positive as at December 1997 (UNAIDS, 1997).

Pattern III

This pattern is prevalent in Eastern Europe, North Africa, the Middle East and Asia, and is characterised by limited transmission and few AIDS cases. For example, 0,13% of adults living in North Africa and the Middle East, and 0,07% of adults living in Eastern Europe and Central Asia were estimated to be HIV-positive as at December 1997 (UNAIDS, 1997).

2.1.6 Impact

A number of authors have attempted to define and quantify the consequences of AIDS (Broomberg *et al*, 1991; Barnett and Blaikie, 1992; Chin *et al*, 1992; Cross, 1993; Whiteside, 1993; UN, 1994; and Patel, 1995). The most devastating consequences are expected to occur in the developing world, and will be briefly discussed at five levels:

Individual Level

Life expectancy will decline. The quality of life for an individual (whether infected or not) living in AIDS-affected areas may decline as poverty and social disruption increases and medical care is redirected.

Household Level

Families of an infected individual may experience financial difficulty, due to the high medical costs, funeral expenses, and possible reduction in family income. Care of the infected family member may involve a great deal of time, and is likely to be physically and emotionally exhausting. If one or more of the breadwinners die, children will be orphaned, often resulting in a lower level of education and higher mortality.

Community and Firm Level

Communities will have to cope with poverty and the increased number of orphans.

Individual firms will experience a loss of trained labour and a reduction in their ability to recruit skilled labour, resulting in a shortage of skills. Costs will rise as productivity declines, and as more resources are needed for training. The cost of morbidity and mortality benefits will increase.

National Level

Certain sectors of the economy are expected to be severely affected by the epidemic. For example, the mining industry in South Africa is expected to lose a substantial portion of their labour force (Foster, 1996). GDP growth may slow, and GDP per capita may well decrease.

Governments will be faced with the problem of resource allocation. A larger health budget with altered priorities may be necessary to cope with the high costs of treating HIV patients, as well as the rapid increase in patient numbers. Funds will also be required for AIDS-awareness campaigns.

If current trends continue, it is estimated that AIDS will cost South Africa one percent of GDP by 2005, and that 75% of the health budget will be used to cover the costs of HIV/AIDS (Mandela, 1997).

International Level

International co-operation will be required in order to curb the spread of the epidemic. Severely affected countries will need additional aid from donor countries.

2.2 The South African Position

2.2.1 The South African Disease

The first South African AIDS case was reported to the health authorities in 1982. From the available data, it appears that two distinct patterns of spread exist in South Africa. Pattern I transmission occurs mainly amongst the white population, and dominated the early spread of the disease. It has since been controlled. Pattern II transmission occurs mainly amongst the black population. This epidemic is believed to have started in the late 1980s (Department of Health, 1996).

Preventative measures appear to have had little effect as yet. A number of reasons may be given:

- Treatment and control of sexually transmitted diseases in South Africa has been "poor" (Schneider, 1993). These diseases act as co-factors in the transmission of HIV.
- Condom usage amongst South African men was found to be extremely low at around 1% (van der Linde, 1996). They were also found to be associated with "filth, disease, infidelity and promiscuity", epitomising the stigma hindering education campaigns.
- Certain black cultures view the practice of multiple sexual partners as acceptable for men (Mandevu, 1995).
- Certain black cultures allow men to choose young girls as sexual partners (Barnett and Blaikie, 1992; Fleming, 1993; and WHO, 1994). These so-called "sugar-daddies" have been known to exchange school fees and other gifts for sexual favours (Chimombo, 1995), and young girls have little power to refuse such advances (Mandevu, 1995).
- The incidence of rape is reported to be high amongst South African women (Lachman, 1995:384).
- A large section of the population is living in poverty. It was estimated in October 1995 that 53% of the population lived in the poorest 40% of households and spent less than R385 per adult per month. (South African Institute of Race Relations, 1997:664). Diseases which suppress the immune system are likely to spread in such conditions, thus increasing the probability of HIV infection. Individuals living in these conditions may also be unable to receive or understand AIDS education campaigns (Whiteside, 1993).

The mean incubation period of HIV infection amongst South African adults is believed to be about 8,5 years (Doyle, 1991). The major opportunistic infections causing death in South African AIDS patients include tuberculosis, bacterial infections, cryptococcal meningitis and PCP (Jentsch, 1996). This is discussed further in section 6.1.

2.2.2 The Current Situation

In 1990, the Department of Health initiated an annual survey amongst pregnant women attending public antenatal clinics throughout South Africa. Blood specimens are anonymously tested for HIV, and the resulting HIV prevalence rates are important in monitoring the spread of the epidemic amongst the sexually active population.

The published 1996 results were found to be inaccurate due to an error in the Western Cape figure. The prevalence rates were later corrected (*AIDS Scan*, 1997). Table 2.1 indicates the corrected national prevalence rate for 1996, as well as the observed prevalence rates for 1990 to 1995 (Department of Health, 1991;

1992; 1993; 1994; 1995 a; and 1996). The rate of increase in infection has dropped from 80% between 1990 and 1991, to 25% between 1995 and 1996. The 1996 figures were used to estimate that 6% of South Africans (or 2,4 million) were infected during 1996 (*Business Day*, 1997). It was also estimated that 11% of males aged 15 to 45 and 14% of females between these ages were HIV-positive during 1996 (*The Star*, 1997).

Table 2.1 Observed national HIV prevalence rates amongst women attending antenatal clinics in South Africa

Year	Prevalence rate
1990	0,89%
1991	1,60%
1992	2,80%
1993	4,90%
1994	8,38%
1995	11,32%
1996	14,17%

Source: Antenatal clinics of the public health service

Provincial prevalence rates are shown in Table 2.2. The provinces with the most advanced epidemics are the North West, KwaZulu-Natal and the Free State. Of concern is the apparent dramatic increase in the North West. Health Minister, Nkosazana Zuma, reasoned that 1994 and 1995 data may have been inaccurate (*Business Day*, 1997). Surveys conducted before 1995 did not include Bophuthatswana in the North West sample, and the progress of the epidemic in this province thus cannot be interpreted over time. KwaZulu-Natal led the national epidemic prior to 1996.

The age distribution of the 1994, 1995 and 1996 prevalence rates is shown in Table 2.3. Prevalence rates have consistently been highest amongst females aged 20-24. The highest rate of increase has also occurred amongst this group. The estimates in the 45-49 age group are based on very small numbers, and are thus unreliable.

AIDS cases are reported to the relevant authorities on a voluntary basis. Figure 2.1 illustrates the age and gender distribution of AIDS cases reported as at 30 November 1995 (Department of Health, 1995 b). Under-reporting of AIDS cases is believed to be considerable. The largest proportion of reported female AIDS cases occur in the age group 20-24, in comparison to age group 30-34 for males.

AIDS-related deaths may or may not be reported to the relevant authorities. Even if reported, the reported cause of death may not relate to AIDS. This point is discussed in greater detail in section 6.1. The implication of this is that AIDS-related deaths are likely to be substantially under-reported. Actual AIDS-related deaths are estimated in Chapter 5.

Table 2.2 Observed provincial HIV prevalence rates amongst women attending antenatal clinics in South Africa

Province	1994	1995	1996
Western Cape	1,16%	1,66%	3,09%
Northern Cape	1,81%	5,34%	6,47%
Eastern Cape	4,52%	6,00%	8,10%
Free State	9,19%	11,03%	17,49%
Gauteng	6,44%	12,03%	15,49%
Mpumalanga	12,16%	16,18%	15,77%
North West	6,71%	8,30%	25,13%
Northern Province	3,04%	4,89%	7,96%
KwaZulu-Natal	14,35%	18,23%	19,90%

Source: Antenatal clinics of the public health service

Table 2.3 Observed HIV prevalence rates amongst women attending antenatal clinics in South Africa, by age

Age group	1994	1995	1996
< 20	6,47%	9,50%	9,50%
20-24	8,94%	13,12%	17,52%
25-29	8,63%	11,03%	15,21%
30-34	6,37%	8,05%	12,13%
35-39	3,72%	7,37%	9,67%
40-44	5,28%	4,36%	9,94%
45-49	0,41%	7,45%	5,83%

Source: Antenatal clinics of the public health service

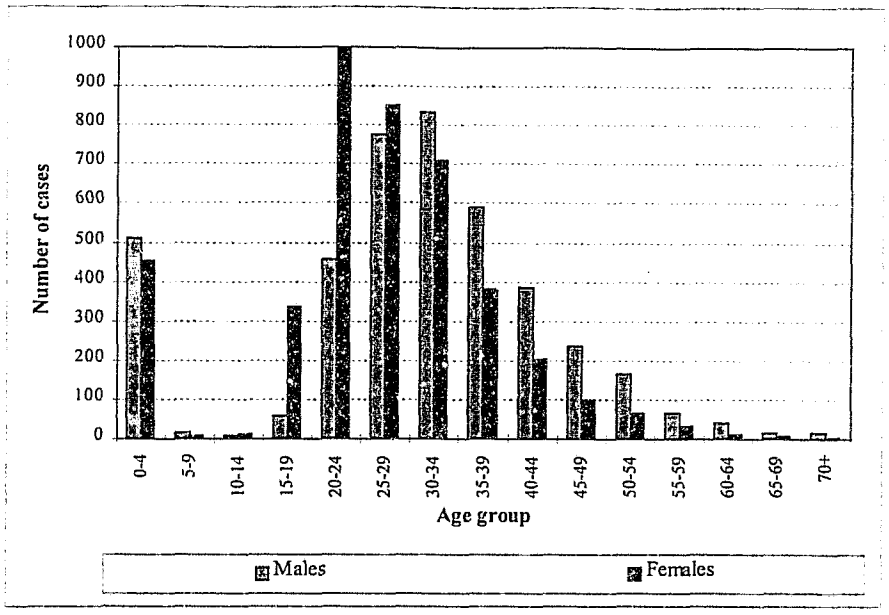


Figure 2.1 The age and gender distribution of reported AIDS cases as at 30 November 1995

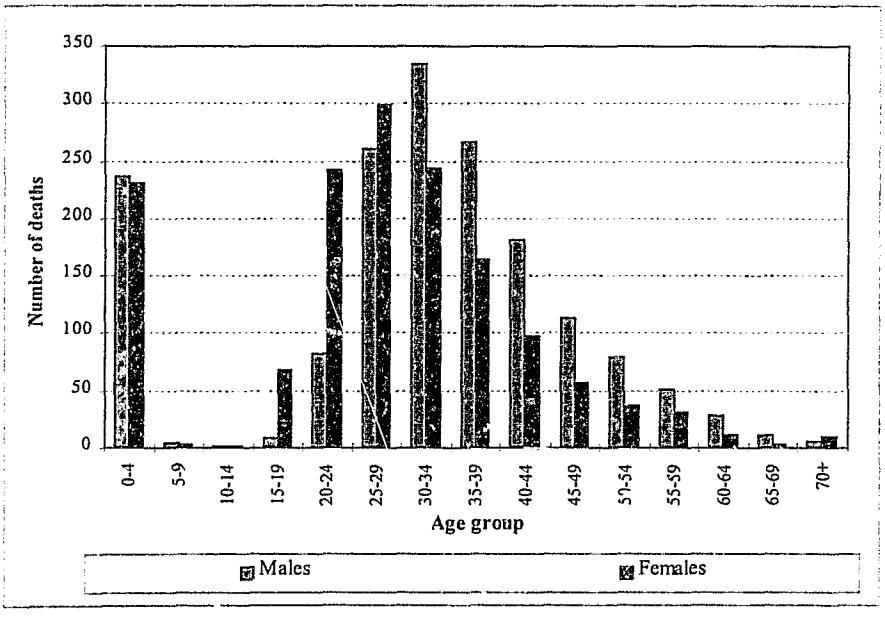


Figure 2.2 The age and gender distribution of AIDS-related deaths reported during 1994

Figure 2.2 illustrates the age and gender distribution of AIDS-related deaths reported during 1994 (CSS, unpublished). The largest number of male deaths occur in age group 30-34, whereas the largest number of female deaths occur in the age group 25-29. This suggests that females become infected some five years younger than males.

3

An Overview of AIDS Models in South Africa

This chapter discusses various models which have been developed to model the spread of the AIDS epidemic. Projection methods are contrasted, and results are reconciled where possible. The reliance upon HIV prevalence data rather than data on AIDS-related deaths is noted.

3.1 Background to Mathematical Modelling

Mathematical modelling is a form of problem-solving involving the transformation of real-world problems into solvable mathematical problems, and translating the mathematical solutions back into real-world solutions.

In order to be mathematically tractable, the real-world problem is usually simplified whilst retaining the essential features of the problem. A trade-off between mathematical tractability and the accurate representation of reality must be made. Models which are over-simplified may not reflect reality, whereas complex models may be costly in terms of computing time.

Real-world problems can often be represented by a number of mathematical models. No model is "incorrect" (Bowie *et al*, 1996), although reasonably simple models which adequately represent reality will be preferable. Nevertheless, even inadequate models may provide insight into certain aspects of the problem.

Extensive literature on mathematical modelling is available, such as Cross and Moscardini (1985), Kapur (1988), and Fowkes and Mahony (1994), as listed in the Bibliography.

3.2 The Structure of AIDS Models

AIDS models can be classified based on their underlying assumptions and

projection methods. Different categories have been suggested by Padayachee and Schall (1990), Doehring (1991), and Doyle *et al* (1997). In this dissertation, AIDS models will be classified into categories according to three criteria.

3.2.1 Criterion 1: Level of Detail

Macro-level models

The epidemiological spread of the disease is modelled at the societal or "macro" level, ignoring the individual events which combine to form the epidemic. Such a model is often a simple extrapolation of an existing trend in HIV prevalence or number of AIDS cases. Macro-level models tend to over-simplify reality by ignoring relevant data, and may not be suitable for long-term projections.

Micro-level models

The behaviour of individuals is modelled in order to determine the spread of the epidemic. Less simplification is involved, and projections are likely to be more accurate. However, micro-level models require large amounts of data which are often unreliable or unavailable, in which case estimates must be made. These estimates introduce inaccuracies into the results. Micro-level models are generally complex and may require extensive computer resources.

3.2.2 Criterion 2: Mathematical Structure

Deterministic models

These models produce exact, mathematically determined results, and thus ignore random error.

Stochastic models

Projections are based on variables which are assumed to follow particular statistical distributions, thus incorporating an allowance for random error. Simulations are often required in order to produce results.

In contrast to deterministic models, stochastic models will generally produce different results if run a number of times with the same assumptions (unless the pseudo-random number generations start at the same point for each run).

3.2 Criterion 3: Calibration

AIDS models may be calibrated to data on:

HIV prevalence, or
AIDS cases, or
AIDS-related deaths,

or a combination of these. Obviously, models fitted to all three will be the most accurate.

The majority of existing AIDS models have been calibrated to HIV prevalence data, rather than data on AIDS cases or AIDS-related deaths. There are a number of reasons for this:

- AIDS cases are extensively underreported, even if identified by health workers. It is estimated that only 10% of AIDS cases are reported in certain African countries, whereas up to 80% may be reported in developed countries (Schoub *et al*, 1990). The South African experience is likely to fall somewhere in-between.
- A large number of AIDS cases are never identified or are incorrectly diagnosed. This is especially relevant in the rural areas where health care facilities are poor.
- The reported number of AIDS cases is understated due to late reporting. An AIDS case may only be identified once the patient requires hospitalisation, or once death has occurred.
- Periodic changes in diagnostic criteria may affect statistics on AIDS cases (Schoub *et al*, 1990).
- AIDS-related deaths may not be recorded as such - i.e. the recorded cause of death may be attributed to some other opportunistic infection (discussed further in section 6.1).
- The South African epidemic is not very mature. Since AIDS-related deaths occur approximately ten years after HIV infection takes place, the number of AIDS-related deaths has been small relative to the number of infected individuals.
- AIDS-related deaths do not give a clear picture of the current state of the HIV epidemic.

3.3 The Model Produced by Edelston

Edelston (1988) developed a deterministic, macro-level model which extrapolates the trend in reported AIDS cases on the basis of some fairly rudimentary assumptions. The number of AIDS cases and HIV-infected individuals are projected for each year up to 2001.

The World Health Organisation (WHO) requires each country's health authority to report diagnosed AIDS cases. The total number of South African AIDS cases reported up to 1 January 1988 was used as the starting point for the projections. There are two sources of under-reporting. Firstly, not all AIDS cases are reported to the South African health authorities. Secondly, AIDS cases which are reported to the health authorities may not be reported to the WHO. These errors may be substantial and are ignored by the model. Projections are therefore likely to understate the true size of the epidemic if all other assumptions are valid.

The model assumes that 104 HIV-infected individuals exist for every reported full-blown AIDS case, as estimated by the WHO. 84 South African AIDS cases had been reported to the WHO by 5 October 1987 (the last date of reporting prior to 1 January 1988). Assuming that the number of reported AIDS cases doubles each year, it was estimated that 99 AIDS cases would have been reported by 1 January 1988. This translates into an additional 10 296 HIV-infected individuals. The size of the South African population was estimated to be 34 855 800 at that date, and was assumed to remain constant.

Edelston produced a number of scenarios based on different assumptions. The two most extreme scenarios are discussed below.

Table 3.1 The number of HIV-infected individuals in South Africa, as projected by Edelston

Date (1 Jan)	Doomsday	Super-Optimist
1988	10 395	10 395
1990	41 580	41 580
1995	1 330 560	398 643
1996	2 661 120	479 976
1997	5 322 240	559 650
1998	10 644 480	636 888
1999	21 288 960	708 855
2000	All	771 225
2001	All	814 422

3.3.1 Doomsday Scenario

It is assumed that no vaccine or cure will be developed before 2001 and that

human behaviour will not change in such a way as to alter the spread of the disease. The number of reported AIDS cases - and hence, the number of HIV-infected individuals - was assumed to double every year. The number of AIDS cases reported to the WHO was taken as a starting point. Results are shown in Table 3.1.

3.3.2 Super-Optimist Scenario

It was assumed that a vaccine was developed in 1994 and that a cure would be found in 1999. The number of new infections was assumed to double each year until 1989, and reduce by one third in each of 1989, 1990 and 1991. A small number of infections were assumed to occur until 1994. The super-optimist scenario is shown in Table 3.1.

3.4 The Model Produced by Padayachee and Schall

Padayachee and Schall (1990) made short-term predictions for the AIDS epidemic in South Africa using three deterministic, macro-level methods.

3.4.1 Extrapolation

This method relied upon HIV prevalence data for blacks, which had been collected from five longitudinal seroprevalence studies:

- Antenatal clinic attendees in the greater Johannesburg area (ANC)
- Blood donors, separately for males (MBD) and females (FBD)
- Attendees of family planning clinics in the greater Johannesburg area (FP)
- Attendees of STD clinics for miners (MIN)
- Attendees of STD clinics in the greater Johannesburg area, separately for males (MSTD) and females (FSTD)

A logistic curve with an asymptote of one was fitted to each of these data sets using standard logistic regression programmes. It was noted that the inferred doubling times were similar for each data set (except FSTD, which was excluded from further analysis), suggesting a mean doubling time of 8,5 months. Each of the curves was refitted with the additional constraint that the doubling time was equal to this mean doubling time. The curves were extrapolated in order to estimate the HIV prevalence amongst blacks aged 15-49 as at 1 January 1990, 1991 and 1992. Projections are shown in Table 3.2.

Table 3.2 HIV prevalence rates for various subgroups of the black population aged 15-49, as projected by Padayachee and Schall

Date (1 Jan)	ANC	MBD	FBD	FP	MIN	MSTD
1990	0,7%	0,4%	0,6%	1,9%	2,3%	2,9%
1991	1,9%	1,1%	1,6%	5,0%	5,9%	7,3%
1992	4,8%	2,8%	4,2%	12,0%	14,0%	17,0%

3.4.2 The Direct Method

The HIV prevalence data for the subgroups listed in section 3.4.1 were used, together with estimates of the population aged 15-49 in each subgroup, to estimate the number of HIV-infected blacks living in the greater Johannesburg area as at 1 January 1989. A high and low estimate of the number of infected black adults in the whole of South Africa was obtained under pessimistic and optimistic assumptions regarding relative levels of infection by location.

Projections for 1 January 1990, 1991 and 1992 were based on these figures, assuming a constant doubling time of 8,5 months. The projections are shown in Table 3.3.

Table 3.3 The number of HIV-infected blacks aged 15-49, as projected by Padayachee and Schall

Date (1 Jan)	Direct Method		Back Calculation	
	Low Estimate	High Estimate	Low Estimate	High Estimate
1990	44 763	63 076	39 984	67 800
1991	119 [illegible]9	167 782	106 318	180 348
1992	316 725	446 300	283 038	479 726

3.4.3 Back Calculation

The number of AIDS cases reported up to 1990 was used, together with an

assumed incubation period distribution, to estimate the number of HIV-infected individuals for years prior to 1990. Low estimates were obtained assuming that all AIDS cases were reported. High estimates were obtained assuming that 59% of AIDS cases were reported, as documented by the US General Accounting Office (USGAO, 1989:21-33).

Forecasts were made assuming a constant doubling time of 8,5 months. The projections are shown together with the results of the direct method in Table 3.3.

The number of HIV-infected black adults projected using the direct method reconcile fairly well with those derived using the method of back calculation. The results thus appear fairly insensitive to the projection method used.

3.5 The Model Produced by Schall

Schall (1990) developed a deterministic model for projecting worst-case scenarios of the South African epidemic amongst the sexually active black population. Assumptions were made about individual behaviour, and the model thus incorporates elements of a micro-level model.

The structure of the model is similar to that of models developed by Bongaarts (1989) and Anderson *et al* (1989). The black population in 1985 (with an estimated size of 25 million) was grouped into cohorts of the same age, gender and sexual behaviour (categorised into four classes). Mortality and fertility rates were used to project the population. Individuals within each cohort were allocated to one of three infection states - susceptible, HIV-infected or AIDS-sick.

Movements between the states were described using sets of linear differential equations. The movement from the susceptible state to the HIV-infected state relied upon assumptions concerning the probability of infection per contact, patterns of partner choice, and the rate of partner change. The movement from the HIV-infected state to the AIDS-sick state was modelled assuming the incubation period followed a gamma distribution with a mean of eight years. Exits from the AIDS-sick state were assumed to follow an exponential distribution, with a mean survival time of one year.

Three worst-case scenarios were projected. Explicitly pessimistic assumptions were made regarding patterns of partner choice and the rate of partner change, and behavioural change and heterogeneity within the black population was ignored. The model thus overstates the magnitude of the epidemic. Parameters were adjusted such that the predicted doubling time at 1 January 1989 was 8,5 months and that the HIV prevalence amongst black women at that date was 0,6%, as in the model produced by Padayachee and Schall (section 3.4).

3.5.1 Scenario 1

In this scenario, it is assumed that sexual partners are chosen at random, and that the rate of partner change is the same for all individuals. These assumptions effectively mean that all individuals are equally at risk of becoming infected, and hence, the epidemic is not contained within certain "high risk" groups.

The results indicate an increase in the size of the black population from 25 million in 1985 to 35 million by the year 2000. The population is then predicted to drop below its initial size by the year 2010, and to reduce to about 12 million by the year 2020. HIV prevalence amongst individuals aged 15 to 49 is estimated to increase from 0% in 1985 to more than 90% by 1995.

3.5.2 Scenario 2

As in Scenario 1, it is assumed that sexual partners are chosen at random. However, different rates of partner change are assumed for each of the four classes of sexual behaviour. The projected epidemic is expected to be less severe than that simulated in Scenario 1, since some allowance is made for the greater promiscuity amongst the "high risk" classes.

The results indicate that the population will increase until some time between 2000 and 2005, and decline slowly thereafter. HIV prevalence amongst individuals aged 15 to 49 is expected to peak between the year 2000 and 2005, at approximately 50% for females and 40% for males.

Schall concluded that Scenario 2 is the most realistic worst-case scenario, as it is very unlikely, although still possible, that the actual epidemic will be as severe. In addition, prevalence rates amongst antenatal clinic attendees in other African countries appear to have peaked at levels between 24% and 35% (Whiteside, 1995).

3.5.3 Scenario 3

In this scenario, it is assumed that two-thirds of sexual contacts will be chosen from the same class of sexual behaviour, and that the remaining one-third will be chosen at random. Different rates of partner change are assumed for the different classes of sexual behaviour, as in Scenario 2. This scenario is the least pessimistic of the three, as the infection is more likely to be contained within the "high risk" classes.

The results indicate that the population will increase until some time between 2005 and 2010, and remain constant at about 40 million thereafter. The HIV prevalence amongst individuals aged 15 to 49 is estimated to peak in the year

2005 at a level of 40% for females and 30% for males.

The author concluded that Scenario 3 is likely to overstate the magnitude of the South African epidemic. However, the assumptions made in this scenario may be realistic, and there is thus a greater chance that the scenario depicts the true size of the epidemic.

3.6 The Model Produced by Doyle

The so-called Doyle model (Doyle and Millar, 1990 and Doyle, 1991) was developed for specific use by Metropolitan Life, and was designed to produce statistics of particular interest to the actuarial profession. These include age, sex and duration specific mortality rates, allowing for both ultimate and select experience. The model is proprietary, and the detailed mathematical structure is not accessible to the public.

The model incorporates micro-level elements, including assumptions regarding risk groups. It is assumed that the population can be categorised into risk groups, each with a similar level of sexual activity. This concept had been applied in previous models (Bongaarts, 1989 and Lörper, 1989), and has recently been used in the model produced by the Actuarial Society of South Africa (described in Chapter 4).

An initial population size, fertility rates and mortality rates are required as inputs to the model. The model can thus be used to project the spread of the AIDS epidemic amongst the total population, or subsets thereof.

The model is deterministic and contains parameters which determine the distribution of the population into the various risk groups, a pattern of sexual contact within and between the risk groups, and rates of HIV infection. Certain parameters can be determined scientifically, whereas others are arbitrary and are thus used for calibration purposes. Allowance is also made for HIV-infected immigrants - a concept developed by Groeneveld (1990). This was seen to be pertinent in the South African context.

The model has been calibrated to HIV prevalence data from the whole of Sub-Saharan Africa, including South Africa. Annual estimates of population size, demographic structure, births, new HIV infections, non-AIDS deaths and AIDS deaths are produced.

A number of scenarios have been produced for the total South African population by making different assumptions about the spread of the epidemic and adjusting the model parameters accordingly. Scenarios have also been produced for certain subsets of the population.

For example, Scenario 60 assumed that sexual behaviour would not change. Scenario 61 assumed that sexual behaviour would alter significantly after a period of 12 years after the start of the epidemic. Scenario 80 is the currently recommended parameter set of the model, although projections have not been published. The Actuarial Society of South Africa has issued a professional guidance note for assured lives based on Scenario 80 (ASSA, 1995).

Scenario 60 indicates that the prevalence amongst the population aged 15 to 49 will peak at approximately 27% by the year 2010. Scenario 61 puts this peak at 18%, occurring at some time between 2000 and 2005. Further estimates from these two scenarios appear in Table 3.4.

Table 3.4 The number of HIV-infected individuals and AIDS-related deaths, as projected by Doyle

Year	Number of HIV-infected individuals		Number of AIDS-related deaths	
	Scenario 60	Scenario 61	Scenario 60	Scenario 61
1991	97 000	97 000	1 350	1 350
1995	970 000	970 000	23 000	23 000
2000	4 112 000	3 700 000	203 000	197 000
2005	6 410 000	4 762 000	525 000	429 000

3.7 The Model Produced by Groeneveld and Padayachee

Groeneveld and Padayachee (1992) developed a stochastic, micro-level model of the epidemic amongst black adults in South Africa. The model is complex and requires a substantial amount of data. Since much of these data were unavailable, data from other countries were used. Best estimates were made, wherever necessary.

The model assumes that HIV-infected immigrants initiated the South African AIDS epidemic. The number of such immigrants is estimated for each year from 1985 to 2000, and ages are assigned according to a normal distribution.

The simulation process begins with the HIV-infected immigrants. Each

individual's sexual activity is modelled, assuming statistical distributions for the number of new sexual partners per year, the frequency of sexual contacts per partner, the probability of infection per contact, and the age of the newly infected partners.

After all infected immigrants have been modelled for the period 1985 to 2000 as described above, the process is repeated for the newly infected individuals who in turn infect other individuals. This process continues until all newly infected individuals have been considered.

In order to estimate the number of AIDS cases in each of the years, the incubation period is assumed to follow a Weibull distribution with a mean of ten years. The model considers each HIV-infected individual separately, and estimates the calendar year in which he/she progresses to full-blown AIDS.

The number of HIV infected individuals and AIDS cases predicted by the model appears in Table 3.5. 29% of the black adult population is predicted to be infected by the year 2000.

Table 3.5 The number of newly HIV-infected individuals and cumulative AIDS cases amongst black adults, as projected by Groeneveld and Padayachee

Date (1 Jan)	Newly HIV-infected individuals	Cumulative number of AIDS cases
1985	415	5
1990	21 684	1 020
1995	297 311	28 215
1996	459 460	52 625
1997	680 565	96 370
1998	961 357	166 362
1999	1 305 211	285 348
2000	1 583 722	499 523

The model is computer intensive and "has not been used beyond the academic setting" (Doyle *et al*, 1997).

3.8 The Model Produced by Gregson *et al*

The model produced by Gregson *et al* (1997) is unpublished, and the underlying mathematical structure has not been clarified.

What is known, however, is that a sigmoid curve has been fitted to HIV-prevalence data for women aged 15 to 49. UN estimates of the population size, fertility rates and mortality rates have been used. Behavioural parameters have been defined so as to fit the observed prevalence trend. Three scenarios have been produced for South Africa based on different assumptions of behaviour change.

The low, intermediate and high scenarios result in a peak HIV-prevalence amongst women aged 15 to 49 of approximately 12% occurring in 1997, 17% occurring in 1999 and 23% occurring in 2005, respectively.

3.9 The Model Produced by Bos and Bulatao

Bos and Bulatao (1992) produced a stochastic, micro-level model of the AIDS epidemic in order to incorporate AIDS mortality into World Bank population projections. It was developed to project the AIDS epidemic in many parts of the world. The South African epidemic can be modelled by choosing suitable parameters.

Three processes are simulated: the transmission of HIV, the development of AIDS amongst HIV-infected individuals, and the progression from AIDS to death. The processes are modelled separately for adults and children.

HIV-transmission is assumed to occur sexually, perinatally, via blood transfusions, and via infected needles.

In order to model sexual transmission, the adult population was divided into four sexual orientation groups: females (assumed to be heterosexual), heterosexual males, homosexual males, and bisexual males. Subgroups of varying sexual behaviour were defined within each group. The calculation of new infections occurring in each subgroup within a particular year is based on a number of assumptions. In particular, assumptions are required for the proportion of potential sexual contacts which are infected, the probability of transmission per contact, the average number of sexual acts per partner, and the number of new partners per annum. The assumptions are allowed to vary by subgroup.

Transmission via blood transfusions and infected needles was modelled in a similar way to sexual transmission. Assumptions are required for the proportion of infected "agents" (needles or blood), the probability of transmission per exposure, and the frequency of such exposure.

The number of infected newborns was calculated as the product of the number of infected women of reproductive age, the general fertility rate, and the probability of transmission from an infected mother to her child. An adjustment factor can be incorporated to allow for higher/lower fertility amongst HIV-infected women.

The progression from HIV-infection to AIDS was assumed to follow a logistic curve. A similar curve was chosen to model the progression from AIDS to death. The parameters were chosen so as to allow a faster progression amongst children.

The model incorporates features which allow the impact of certain variations to be measured:

- A proportion of individuals may be assumed to move randomly between subgroups.
- Certain parameters may be varied over time. For example, rates of infection may vary by duration since infection.
- Infection rates may be adjusted to allow for cofactors (e.g. condom use, the presence of genital ulcers, etc.).
- The adult population can be divided into groups of differing susceptibility to infection. Genetic variability of individuals can thus be incorporated.
- The likelihood of sexual contact between each pair of subgroups can be altered. Urban-rural differences can be modelled by defining separate urban and rural subgroups, and limiting sexual contact between the subgroups.

The model is linked to the World Bank's standard demographic model (Hill, 1990) which was used to produce population projections in the absence of AIDS.

A major obstacle to running the model for Sub-Saharan Africa is the unavailability of inputs. Bos and Bulatao made assumptions to simplify the model (e.g. homosexual and bisexual transmission was ignored). Simulations were performed separately for each country in the region, and unknown parameters were varied until estimated actual levels of HIV prevalence were achieved.

The population was projected up to the year 2050. The difference in the total size of the South African population under an AIDS and no-AIDS scenario was projected to be 0,1% in 2000, 0,6% in 2020, and 0,8% in 2050. This is significantly less than the 2,6%, 10,0% and 14,9% projected for Uganda in the three respective years.

3.10 The Model Produced by Chin and Lwanga

Chin and Lwanga (1991) of the World Health Organisation developed a simple, macro-level model for the estimation and short-term projection of adult AIDS cases. A computer program (referred to as the "Epi Model") was developed to enable users to run the model with different parameter values. The model can be

used for different countries by choosing suitable parameters.

A fundamental assumption of the model is that cumulative HIV infections (excluding AIDS cases) follow a sigmoid curve. Furthermore, it is assumed that HIV incidence over time follows a gamma distribution. The parameters were chosen so as to provide the best fit to reported AIDS cases in countries with reliable reporting systems. An assumption is required for the year during which extensive spread of HIV started. "Extensive spread" was defined as an HIV prevalence of at least 1% within a high-risk group.

The cumulative HIV infection curve is defined using a point HIV prevalence figure (excluding AIDS cases). This point prevalence underestimates the cumulative HIV incidence, since some infected individuals would have developed AIDS and perhaps died. The model adjusts the point prevalence upwards, using annual progression rates from HIV to AIDS. These progression rates were obtained from published cohort studies conducted by the WHO. The cumulative HIV incidence is set equal to this adjusted prevalence.

An assumption is required regarding the position (called the reference year) of the point prevalence on the HIV incidence curve. As a starting point, it was assumed that the reference year coincided with the year of peak incidence. It was found that moving the reference year to three years either side of the year of peak incidence made little difference to projected AIDS cases.

The cumulative HIV incidence is divided into annual cohorts of HIV infections using the gamma distribution. The cumulative number of AIDS cases is projected using the progression rates from HIV to AIDS.

The model was used to project the epidemic for Sub-Saharan Africa as a whole. A point prevalence of 5 million HIV-infected adults in 1990 was used, and 1992 was assumed to be the year of peak incidence. It was estimated that there were close to 700 000 cumulative AIDS cases by the end of 1990. This was projected to reach the 2 million mark by 1994.

3.11 The Model Produced by Bongaarts

Bongaarts (1989) produced a model under the auspices of the Population Council to project the annual incidence and prevalence of HIV and AIDS. The model is micro-level, as it examines interaction between individuals.

The spread of the disease was modelled on top of a standard demographic framework. The underlying population was divided into cohorts based on year of birth. Standard demographic assumptions were required to model the dynamics of the population. Each cohort was divided further into subgroups of varying sexual activity based on assumptions about the population. Individuals were assigned an

infection status: infected or uninfected. The uninfected state was assumed to be made up of two parts: those susceptible to HIV infection, and those immune to infection. However, Bongaarts acknowledged that studies had not been able to confirm the existence of the latter genetic trait.

Transmission was modelled using sets of linear differential equations. Two routes of transmission were modelled: sexual (heterosexual, homosexual and bisexual), and perinatal.

The probability of infection was assumed to be a function of the frequency of contact, the probability of choosing an infected partner, the infectiousness of an infected partner, and the rate of partner change. Infectiousness was assumed to depend on gender, stage of disease, the presence of any cofactors, and condom usage. The sexual contact rate was assumed to vary with stage of disease.

HIV infected individuals were assigned to two groups: those at risk of developing AIDS, and those who are not at risk of developing AIDS. For those at risk, the progression from infection to AIDS was assumed to follow a gamma distribution. The parameters were chosen so as to correspond to a mean incubation period of 9,3 years. An assumption was made regarding the monthly mortality rate amongst individuals with AIDS.

The entire model was simulated as a Markov chain, which made for fast implementation on a computer.

Bongaarts applied the model to a Central African country by choosing suitable demographic assumptions. All sexual contact was assumed to be heterosexual. The projected population under the AIDS scenario was 10,8% lower than that projected under the no-AIDS scenario after 25 years. Population growth was projected to decline, but would not become negative due to the high fertility assumptions. HIV prevalence amongst adults was projected to reach 21% after 25 years. The ratio of female to male infected adults increased from 1,0 in the early years to 1,15 in year 25.

3.12 Other Models Produced Overseas

The models discussed in sections 3.9, 3.10 and 3.11 are well known and used by influential organisations. However, a large number of less widely used models have also been produced overseas. Among these are Isham, 1988; Brookmeyer and Damiano, 1989; Tan and Hsu, 1989; Wiley, Herschkorn and Padian, 1989; Zeger, See and Diggle, 1989; Golub, Gorr and Gould, 1993; Rusenberg, Cooke and Hsien, 1995; and Thomas, 1996, as listed in the Bibliography. Most of these rely on fairly complex mathematical derivations. None have been applied directly to the South African epidemic.

3.13 Discussion and Comparison

Edelston's doomsday scenario has not been realised in practice. Doubling times have reduced to less than one year and the projected prevalence rate of 60% in 1999 is not feasible.

The methodology used in Edelston's model has been sharply criticised in an editorial in the South African Medical Journal (Schall and Padayachee, 1990). The use of this method for long-term forecasts was described as "simplistic and unscientific". Further, the authors warned that such sensational doomsday forecasts could result in apathy, if believed, or may damage the credibility of future realistic AIDS warnings, if disbelieved.

Edelston's super-optimistic scenario reconciles with Padayachee and Schall's estimates for 1990. However, the number of infected individuals is understated by 570 000 in 1995, and by approximately 3 million in 2000 when compared to Doyle's estimates.

Padayachee and Schall's use of the logistic curve may be criticised, as HIV-prevalence is unlikely to tend to 100%. Moreover, an exponential curve would probably be sufficient for short-term projections.

Their estimates of infected individuals in 1991 (between 40 000 and 63 000) are slightly lower than Doyle's estimate of 97 000 under Scenarios 60 and 61. They also suggest a greater number of infections in 1990 and 1991 than projected by Groeneveld and Padayachee. The large annual increase may be a result of the assumption of a constant doubling time.

Schall's Scenario 1 has not been realised in practice. Such prevalence rates have not been observed even in the "high risk" urban areas of Central Africa, and are thus unrealistic for South Africa, even for the projection of the maximum size of the epidemic.

Schall's Scenarios 2 and 3 appear unrealistically pessimistic when compared to Doyle's estimates. Conversely, the estimates produced by Gregson *et al* seem to indicate a lower peak prevalence than Doyle's projections.

Groeneveld and Padayachee's prediction that 29% of the black population will be infected by the year 2000 accords with Schall's Scenario 3. The predictions may thus be seen as worst-case estimates. Although the model is stochastic and involves simulation, no indication is given of the variability of the results.

Doyle's model has been the most widely used and accepted model of the South African AIDS epidemic.

Bos and Bulatao estimated that AIDS would result in a South African population

which was 0,6% smaller than the level it would otherwise have reached in 2020. The statistic was estimated to be 10,0% for the Ugandan population. This is more conservative than Bongaarts' estimate of 10,8% for a Central African country.

The results of the model produced by Chin and Lwanga cannot be compared, as it was applied to Sub-Saharan Africa as a whole. A frequent criticism of this model's structure is that many curves can be drawn through a single point (the point HIV prevalence). The advantage of such a macro-level model, however, is that few inputs were required. This was especially relevant when the model was produced (1991), as little empirical evidence was available.

3.14 Summary

A useful model must satisfy two constraints: it should be detailed enough to provide accurate predictions, whilst at the same time being a simplification of reality. Macro-level models are not useful in the long run as their predictions are likely to become less accurate over time. Conversely, models which involve a large number of parameters are not useful if the parameters cannot be reasonably estimated.

Edelston's model is unrealistic due to the assumption of a constant doubling time. Padayachee and Schall's estimates were also based on this assumption, but did not attempt to project into the long-term.

Schall's model used linear differential equations to project three worst case scenarios, rather than the true size of the epidemic. The mathematical structure is very similar to that of Bongaarts' model.

The mathematical structure of Doyle's model cannot be examined or adjusted. The model has been highly reputed as results are believed to be accurate. Moreover, the actuarial profession has based mortality recommendations on these projections.

Groeneveld and Padayachee's model is very complex, as it models the spread of the epidemic at the individual level. Extensive computing resources are required to run the model.

The model pr[illegible]d by Gregson *et al* has not yet been completed.

The overseas models were not specifically produced to model the South African epidemic. However, by choosing parameters relevant to the South African situation, the impact of the South African epidemic can be projected.

Bongaarts' model uses linear differential equations to model transmission of the virus. The model was simulated as a Markov chain and implemented on computer.

Chin and Lwanga produced a macro-level model which fits a sigmoid curve to a point HIV prevalence. It was intended to be used for short-term projections only. The production of "Epi Model" allowed users to change parameters as required.

Bos and Bulatao produced a micro-level model which incorporated all forms of transmission. A number of simplifications were required when applying to Sub-Saharan Africa due to the lack of empirical data.

None of the models discussed in this chapter are suitable for calibration. The model which will be calibrated in this dissertation is described in Chapter 4.

4

The ASSA Model in Detail

This chapter focuses on the model built by the Actuarial Society of South Africa (referred to as the ASSA model), and elaborates on its assumptions, inputs, projection methods and results.

4.1 Background

Life assurers set current premium rates and reserves at a level based on conservative estimates of the future. They are based on projections of, amongst other factors, mortality rates, permanent and total disability rates, rates of sickness and loss of man-hours at work. The AIDS epidemic is likely to have a significant - although uncertain - impact on each of these factors.

Insufficient premium rates or reserves may lead to large unexpected losses. Premiums which are too large may reduce competitiveness. Over-cautious reserves will lead to lower declared profits and could reduce competitiveness. The actuaries, who actually set the level of premium rates and reserves, thus carry a responsibility, and in response have produced a number of models which attempt to project the future impact of AIDS as accurately as possible.

Actuaries look to their professional body, the Actuarial Society of South Africa (ASSA), for guidance. The ASSA has produced an AIDS model for the recommendation of general guidelines. Individual life offices can, however, make their own adjustments.

The model was built by the ASSA AIDS Committee, and was released in September 1996. It is available, free of charge, on the Internet at web site:
http://www.und.ac.za/und/eco/eru/eru.htm

4.2 An Overview of the Structure of the Model

The ASSA Model has been implemented in Microsoft Excel. The programs (macros) have been written in Visual Basic.

The model projects the spread of the AIDS epidemic amongst the black population in South Africa, assuming that HIV was introduced to this population in 1985. The virus is assumed to be transmitted heterosexually, as well as vertically. All other forms of transmission (homosexual, between intravenous drug-users, via blood transfusions) are ignored. The model is thus of the so-called "Pattern II" epidemic, which is the dominant pattern of spread in Africa.

The economically active population (individuals aged between 14 and 60) is split into four groups, based on the risk of becoming infected with HIV via heterosexual contact. This risk is dependent on the level of sexual activity. It is assumed that individuals do not move between risk groups.

The risk groups are described below:

- PRO: Individuals characterised by a very high level of promiscuity, such as prostitutes and their frequent clients.
- STD: Individuals characterised by a high prevalence of sexually-transmitted diseases (STDs). Note that the presence of a STD increases the risk of HIV infection per contact, and is suggestive of a relatively large number of new sexual contacts per annum.
- RSK: Individuals characterised by stable sexual relationships, but who are at risk since they, or their partners, have had one or more other sexual contacts in the past.
- NOT: Individuals who are not at risk of HIV infection, either because they have had a long-term monogamous relationship, or because they have had no sexual contact.

Once individuals reach the age of 60, they are allocated to a common category (OLD) regardless of their previous risk status. The model assumes that no further infections take place amongst individuals in this category, and that all individuals die by the age of 90.

Children younger than 14 are grouped in a YOUNG category, and it is assumed that no sexual transmission of HIV occurs at these ages. This is unrealistic if significant AIDS-related deaths occur amongst teenagers. The model incorporates vertical transmission, and infants may thus be infected at birth.

The population $P(t)$ at point t in time is thus broken up into a finite number of cells based on the values of various dimensions (gender (s), age (x), risk group (r), HIV status (c) and duration since infection (d)). $P(t)$ is a vector with elements $P_i(t)$ denoting the number of individuals in state i. Each state i is defined by a specific

set of values of the dimensions s, x, r, c and d, and $P_i(t)$ can therefore alternatively be denoted $P_{(s,x,r,c,d)}(t)$. Different states differ with respect to at least one of these dimensions.

The model is essentially a multistate model satisfying equation 4.1:

$$P(t+1) = M(t).P(t) \qquad (4.1)$$

where $M(t)$ is a movement matrix with element $m_{ij}(t)$ representing the probability of moving from state i to state j between time t and time $t+1$. A number of these elements will be zero (e.g. the probability of moving from one gender to another, the probability of moving from age x to x-1, etc.). The movement matrix does not incorporate an error term to account for the random nature of the spread of the disease, and the parameters do not follow statistical distributions. The model is thus deterministic (rather than stochastic).

Note that since $P(t+1)$ is dependent only on $P(t)$ (and not the population at any previous time), the model is a Markov chain (as are life tables and the death/sickness models, familiar to those working in the life assurance industry).

A number of HIV-positive immigrants are assumed to start the epidemic in 1985 through sexual contact with the economically active population. These immigrants are assumed to be sex workers, and thus exhibit the sexual activity of the PRO risk group.

4.3 Inputs

4.3.1 Notation

The following notation is required:

Dimensions

- Let $t = 0, 1, 2, ..., 20$ denote 1 January 1985+t. The calendar year from t to $t+1$ will be referred to as year t.
- Let $s = 0, 1$ denote males and females respectively.
- Let $x = 0, 1, 2, ..., 90$ denote age.
- Let $r = 1, 2, 3, ..., 6$ denote the risk groups YOUNG, PRO, STD, RSK, NOT and OLD respectively.
- Let $d = 0, 1, 2, ..., 19$ denote curtate duration since HIV infection.

Not all combinations of s, x, r and d are valid, and the following restrictions apply:

a) $r = 1 \Leftrightarrow x \leq 13 \Rightarrow d = x$ (if HIV-positive)
b) $r = 2, 3, 4, 5 \Leftrightarrow 14 \leq x \leq 59$
c) $r = 6 \Leftrightarrow x \geq 60 \Rightarrow d \leq x - 60$ (if HIV-positive)

Inputs

- Let $P(0)$ denote the total size of the starting population on 1 January 1985.
- Let p_x denote the proportion of the starting population aged x last birthday for $x \neq 90$, and the proportion of the population aged x and over for $x = 90$.
- Let σ_x denote the proportion of males constituting the starting population aged x last birthday for $x \neq 90$, and the proportion of males constituting the population aged x and over for $x = 90$.
- Let g_r denote the proportion of the population aged between 14 and 60 who are in risk group r. Note that $g_1 = g_6 = 0$, and $\sum_{r=2}^{5} g_r = 1$.
- Let I_s denote the number of HIV-infected immigrants with gender s in each year t.
- Let $j_{s,x}$ denote the proportion of HIV-infected immigrants with gender s who are aged x last birthday for $x < 90$, and who are aged x and over for $x = 90$. Note that $\sum_{x=0}^{90} j_{s,x} = 1$.
- Let TFR_r denote the total fertility rate for risk group r.
- Let t_x denote the proportion of the total fertility rate for the total population which applies to women aged x last birthday.
- Let γ_s denote the proportion of babies born with gender s. Clearly, $\gamma_0 + \gamma_1 = 1$.
- Let $q_{s,x}$ denote the non-AIDS rate of mortality for an individual with gender s and aged x last birthday.
- Let β denote the proportion of babies born to HIV-positive mothers who become infected at birth.
- Let θ denote the rate of mortality due to AIDS amongst HIV-positive individuals in the YOUNG risk group.
- Let $i_{s,x}$ denote the value of the sexual activity index for individuals with gender s and aged x last birthday.
- Let Ω_{r_i,r_j} denote the average number of individuals in risk group r_j that an infected individual in risk group r_i would be expected to infect in one year, given that he/she only had contact with uninfected individuals.
- Let λ denote the relative infectivity factor of males to females.
- Let s_d denote the proportion of HIV-infected individuals surviving d full years after infection.
- Let m denote the median term to death of an HIV-positive individual (in years).

4.3.2 Starting Population

The model requires an assumption regarding the size of the black population as at 1 January 1985 to project the population size at future points in time. All individuals are assumed to be HIV-negative at time $t = 0$. This population is then distributed by gender, age and risk group using inputted values of p_x, σ_x and g_r. The number of individuals with gender s, aged x last birthday and in risk group r (denoted $C_{s,x,r}(0)$) is calculated as follows:

$$\begin{aligned} C_{s,x,r}(0) &= P(0).p_{s,x}.g_r && \text{for } 14 \leq x \leq 59 \\ C_{s,x,r}(0) &= P(0).p_{s,x} && \text{for } x < 14 \text{ or } x > 59 \end{aligned} \qquad (4.2)$$

$$\begin{aligned} \text{where } p_{s,x} &= p_x.\sigma_x && \text{for } s = 0 \\ \text{and } \quad p_{s,x} &= p_x.(1-\sigma_x) && \text{for } s = 1 \end{aligned}$$

$P(0)$ was set equal to 25 million, and p_x and σ_x accord with the structure of the black population estimated by the 1994 October Household Survey (CSS, 1995). The assumed values of g_r are 4%, 28,5%, 37,5% and 30% for the PRO, STD, RSK and NOT risk groups respectively. The values were defined in conjunction with values of Ω_{r_i,r_j}, as discussed in section 4.3.7.

4.3.3 HIV-Infected Immigrants

I_0 male and I_1 female infected immigrants are assumed to live in South Africa during each year between 1985 and 2005. These immigrants are assumed to be responsible for introducing the HIV epidemic into South Africa in 1985, and are not included in the South African population. These immigrants are distributed across the individual ages using inputted values of $j_{s,x}$. The number of HIV-infected immigrants with gender s and aged x last birthday (denoted $I_{s,x}$) in each year t is calculated as follows:

$$I_{s,x} = I_s.j_{s,x} \qquad (4.3)$$

I_0 and I_1 were set equal to 120, and values of $j_{s,x}$ appear to have been arbitrarily determined.

4.3.4 Fertility

The model assumes that women will only fall pregnant between the ages of 15 and 50. Total fertility rates are required for each risk group 2 to 5 (denoted TFR_r). The total population fertility rate (denoted TFR) is then calculated as:

$$TFR = \sum_{r=2}^{5} TFR_r . g_r \qquad (4.4)$$

Scaling factors t_x are used to calculate the fertility rates applicable to women aged x last birthday and in risk group r (denoted $f_{x,r}$) as follows:

$$f_{x,r} = TFR_r . t_x \qquad (4.5)$$

The assumed total fertility rates were 1,50, 2,75, 4,50 and 5,25 for the PRO, STD, RSK and NOT risk groups respectively. Equation 4.4 yields an overall total fertility rate of 4,11.

Note that total fertility rates increase as the level of risk decreases. This may appear counter-intuitive as a higher number of children is obviously the result of greater sexual activity. However, women who choose to have children are generally in more stable sexual relationships. In addition, women who choose a promiscuous lifestyle cannot maintain this lifestyle if they fall pregnant, and therefore tend to use contraception.

The t_x factors correspond roughly with the age distribution of the fertility rates used by the CSS to adjust the 1991 census (CSS, 1991 b).

The births resulting from these fertility rates are split by gender according to an assumed proportion γ_s. γ_0 was set equal to 52%.

4.3.5 Non-AIDS Mortality Rates

Non-AIDS deaths occur amongst HIV-negative and HIV-positive individuals. The ASSA model assumes that the same non-AIDS mortality is experienced by HIV-negative and HIV-positive individuals, and that this mortality is independent of t and r. Values of $q_{s,x}$ were obtained from the mortality rates produced by Dorrington, Martens and Slawski (1991). These rates are discussed in section 5.6.1.

4.3.6 AIDS Mortality Rates

Individuals infected via vertical transmission are assumed to experience a constant AIDS mortality rate each year (denoted θ), and are assumed to die before reaching age 14. θ was set equal to 30%.

AIDS mortality rates for HIV-positive individuals aged 14 and over are derived from a table of survival probabilities indexed by curtate duration since infection (s_d). These survival rates follow an arbitrarily determined reverse S-shaped curve with $s_0 = 1$, defined as follows:

$$s_d = \frac{exp(-k(d-m))}{1+exp(-k(d-m))} \qquad \text{for } d \neq 0 \qquad (4.6)$$

where k and m are parameters, and m is the median term to death of an HIV-infected individual (in years).

The AIDS mortality rates are assumed to be independent of s, x, r, and t, and are denoted $q_d{}^A$. They are calculated from the survival probabilities as follows:

$$q_d{}^A = \frac{s_d - s_{d+1}}{s_d} \qquad (4.7)$$

m was set equal to 10 for the default Scenario 500 of the ASSA model. Other scenarios (501 and 502) have been produced by changing this parameter. k was set equal to 0,5.

4.3.7 Infection

The model assumes that a constant proportion (β) of babies born to HIV-infected mothers will be infected. This proportion was set equal to 25%.

Heterosexual HIV infection is assumed to occur amongst individuals aged between 14 and 60 only (i.e. individuals in risk groups 2, 3, 4 and 5). A number of inputs are used to calculate the probability of HIV infection in each year t.

Table 4.1 Values of Ω_{r_i,r_j} assumed in the ASSA model

Risk group r_i	Risk group r_j			
	PRO	STD	RSK	NOT
PRO	1	0,1	0	0
STD	0,1	0,1	0,02	0
RSK	0	0,02	0,01	0
NOT	0	0	0	0

The values assumed under Scenario 500 of the ASSA model for Ω_{r_i,r_j} are reproduced in Table 4.1. Note that the table is symmetric (i.e. $\Omega_{r_i,r_j} = \Omega_{r_j,r_i}$), and

that $\Omega_{r_i,5} = \Omega_{5,r_j} = 0$.

The rationale behind the values in the above table was described by Whitelock-Jones (personal communication):

When HIV first became evident, the doubling time was less than one year. Assuming that the epidemic was initially spread by individuals in the PRO group, it implies that individuals within this risk group infect more than one individual per annum. This is reflected in the table.

As the epidemic has matured, the doubling time has increased. The rate of this decrease was used, together with assumed proportions of the population in the PRO and STD groups (denoted g_2 and g_3), to set the values of $\Omega_{2,2}$ and $\Omega_{2,3}$. The values were adjusted in order to reconcile the projected HIV prevalence with that of sex workers and STD clinic attendees at the early stages of the epidemic. The values of $\Omega_{2,4}$ and $\Omega_{3,4}$ were derived fairly arbitrarily.

The main determinant of the eventual size of the epidemic is the proportion of the population in the NOT risk group (denoted g_5). This proportion was set assuming that the peak HIV prevalence rate in South Africa will be similar to that experienced by other African countries.

The model was calibrated to the HIV prevalence of antenatal clinic attendees, and the values of Ω_{r_i,r_j} and g_r were adjusted to ensure a close fit.

$i_{s,x}$ denotes the value of a sexual activity index for individuals with gender s and aged x last birthday. This index incorporates the relative number of new sexual partners per annum, the relative number of contacts per partner, and the ease of infection per contact. It is independent of risk group for computational ease, although some correlation may exist. The indices are defined by curves of the form:

$$i_{s,x} = k_s(x - a_s)\exp(-b_s(x - a_s)^2) \qquad (4.8)$$

where a_s and b_s are parameters, and k_s is defined such that the mean of $i_{s,x}$ equals one. A value of $i_{s,x}$ greater than one thus indicates higher than average sexual activity, and vice versa.

Females are assumed to be sexually active from age 14 to 50, and males from 16 to 59. Values of a_0, b_0, a_1 and b_1 were defined as 15, 0,004, 10 and 0,005 respectively under Scenario 500 of the ASSA model. Figure 4.1 illustrates these curves.

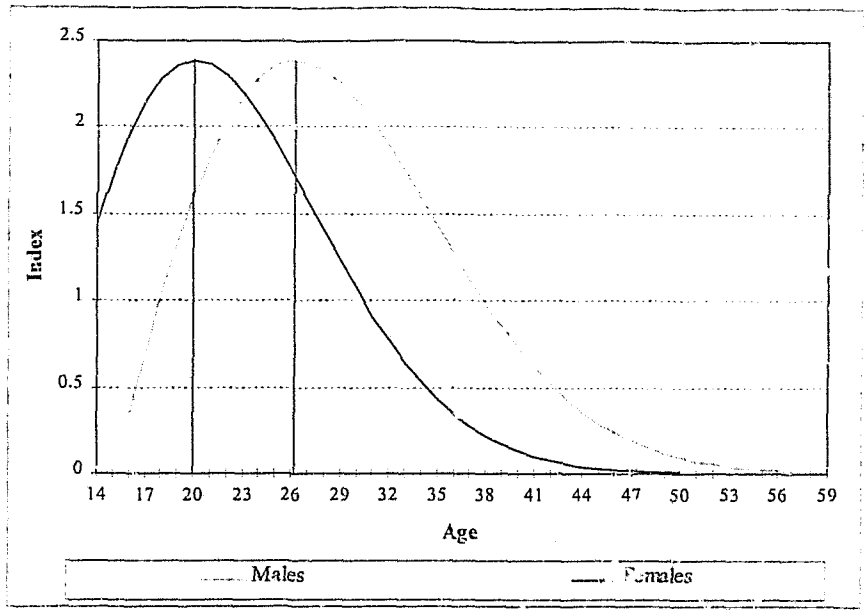


Figure 4.1 Sexual activity indices assumed in the ASSA model

Although the shapes of the sexual activity curves are arbitrary, they do reflect the fact that female sexual activity tends to peak at an earlier age than male sexual activity. This can be corroborated from Figure 2.2, which indicates a higher age at death for males than for females.

λ denotes the relative infectivity factor of males to females. λ was set equal to two in the model, thus implying that females are twice as likely to be infected by an HIV-positive male per contact than vice versa.

4.4 Projections

The following statistics are projected for each t:

- The number of HIV-negative individuals at time t+1, by gender, age and risk group.
- The number of HIV-positive individuals at time t+1, by gender, age, risk group and curtate duration since infection.
- The number of non-AIDS deaths occurring during year t.
- The number of AIDS-related deaths occurring during year t.

In order to calculate these statistics, probabilities of HIV infection are required. As

the epidemic spreads, the probability of infection changes. The model thus calculates a different probability of infection for each year t.

4.4.1 Probability of Infection

A number of quantities need to be defined in order to calculate the probability of infection:

Sexual Activity Weighted HIV Prevalence

Let $C_{s,x,r}(t)$ denote the number of HIV-negative individuals with gender s, age x last birthday and in risk group r at time t, and
let $F_{s,x,r,d}(t)$ denote the number of HIV-positive individuals with gender s, age x last birthday, in risk group r and with curtate duration since infection d at time t.

Let $SP_{s,r}(t)$ denote the "sexual activity weighted HIV prevalence" of an individual with gender s and in risk group r for year t, as follows:

$$SP_{s,r}(t) = \frac{\sum_{x=14}^{59}[i_{s,x}(\sum_{d=0}^{19}F_{s,x,r,d}(t) + I_{s,x})]}{\sum_{x=14}^{59}[C_{s,x,r}(t) + \sum_{d=0}^{19}(F_{s,x,r,d}(t) + I_{s,x}(t))]} \qquad \text{for } r = 2$$

$$SP_{s,r}(t) = \frac{\sum_{x=14}^{59}[i_{s,x}\sum_{d=0}^{19}F_{s,x,r,d}(t)]}{\sum_{x=14}^{59}[C_{s,x,r}(t) + \sum_{d=0}^{19}F_{s,x,r,d}(t)]} \qquad \text{for } r = 3, 4, 5 \qquad (4.9)$$

Partners' Force of Infection

I et $\Delta_{s,r}(t)$ denote the force of infection of all possible partners of an individual with gender s and in risk group r operating during year t, as follows:

$$\Delta_{s,r}(t) = \sum_{j=2}^{5}\Omega_{j,r}.SP_{s',j}(t) \qquad (4.10)$$

where s' denotes the complement of s.

Force of Infection

Let $\delta_{s,x,r}(t)$ denote the force of infection operating during year t on an individual with gender s, aged x last birthday and in risk group r. This force of infection is calculated as follows:

$$\delta_{s,x,r}(t) = 2.i_{s,x}.(\frac{1}{1+\lambda}).\Delta_{s,r}(t) \qquad \text{for } s = 0$$

$$\delta_{s,x,r}(t) = 2.i_{s,x}.(\frac{\lambda}{1+\lambda}).\Delta_{s,r}(t) \qquad \text{for } s = 1 \qquad \textit{(4.11)}$$

The adjustments $2.(\frac{\lambda}{1+\lambda})$ and $2.(\frac{1}{1+\lambda})$ in equation 4.11 effectively multiply the values in Table 4.1 by 4/3 for male-to-female infection, and by 2/3 for female-to-male infection. Gender differences in infection rates are thus allowed for.

Probability of Infection

The probability that an individual with gender s, aged x last birthday and in risk group r will become HIV-infected via heterosexual contact during year t (denoted $w_{s,x,r}(t)$) is calculated as follows:

$$w_{s,x,r}(t) = 1 - exp(-\delta_{s,x,r}(t)) \qquad \textit{(4.12)}$$

Equation 4.12 is analogous to equation 1.6.8 in Neill (1977:16).

4.4.2 HIV-Negative Individuals

The number of HIV-negative individuals with gender s, aged x last birthday and in risk group r at time $t+1$ (denoted $C_{s,x,r}(t+1)$) is calculated as follows:

a) $C_{s,x,r}(t+1) = \gamma_s.\{\sum_{x=14}^{59}\sum_{j=2}^{5} f_{x,j}[\,C_{1,x,j}(t) + (1-\beta)\sum_{d=0}^{19} F_{1,x,j,d}(t)]\}$

for $x = 0$ (and thus $r = 1$)

b) $C_{s,x,r}(t+1) = C_{s,x-1,r}(t).(1-q_{s,x-1})$

for $x = 1, 2, 3, ..., 13$ (and thus $r = 1$) or $x = 61, 62, 63, ..., 90$ (and thus $r = 6$)

c) $C_{s,x,r}(t+1) = g_r.\{C_{s,x-1,1}(t).(1-q_{s,x-1}) + F_{s,x-1,1,x-1}(t).(1-q_{s,x-1}).(1-q_{x-1}^{A})\}$

for $x = 14$ (and thus $r = 2, 3, 4, 5$)

d) $C_{s,x,r}(t+1) = C_{s,x-1,r}(t).(1-q_{s,x-1}).(1-w_{s,x-1,r}(t))$
for $x = 15, 16, 17, ..., 59$ (and thus $r = 2, 3, 4, 5$)

e) $C_{s,x,r}(t+1) = \sum_{j=2}^{5}[C_{s,x-1,j}(t).(1-q_{s,x-1}).(1-w_{s,x-1,j}(t))]$
for $x = 60$ (and thus $r = 6$) *(4.13)*

Note that all individuals aged 13 at time t who survive to age 14 at time $t+1$ (whether HIV-negative or HIV-positive) are assumed to be HIV-negative at time $t+1$. In other words, HIV-positive babies are assumed to die before reaching age 14 or become HIV-negative at age 14.

Note also that all HIV-positive individuals aged x last birthday (where $x < 14$) are assumed to have been infected at birth, and their curtate duration since infection is thus equal to x.

4.4.3 HIV-Positive Individuals

The number of individuals with gender s, aged x last birthday and in risk group r at time $t+1$ who became infected during year t (denoted $F_{s,x,r,0}(t)$) is calculated as follows:

a) $F_{s,x,r,0}(t+1) = \gamma_s.\beta.\sum_{x=14}^{59}\sum_{j=2}^{5}(f_{x,j}\sum_{d=0}^{19}F_{1,x,j,d}(t))$
for $x = 0$ (and thus $r = 1$)

b) $F_{s,x,r,0}(t+1) = 0$
for $x = 1, 2, 3, ..., 14$ (and thus $r = 1$) and $x = 61, 62, 63, ..., 90$ (and thus $r = 6$)

c) $F_{s,x,r,0}(t+1) = C_{s,x-1,r}(t).(1-q_{s,x-1}).w_{s,x-1,r}(t)$
for $x = 15, 16, 17, ..., 59$ (and thus $r = 2, 3, 4, 5$)

d) $F_{s,x,r,0}(t+1) = \sum_{j=2}^{5}[C_{s,x-1,j}(t).(1-q_{s,x-1}).w_{s,x-1,j}(t)]$
for $x = 60$ (and thus $r = 6$) *(4.14)*

The number of HIV-positive individuals with gender s, aged x last birthday, in risk group r and with curtate duration since infection equal to d (where $d > 0$) at time $t+1$ (denoted $F_{s,x,r,d}(t+1)$) is calculated as follows:

a) $F_{s,x,r,d}(t+1) = 0$
for $x = 0$ (and thus $r = 1$) and $x = 14$ (and thus $r = 2, 3, 4, 5$)

b) $F_{s,x,r,x}(t+1) = F_{s,x-1,r,x-1}(t).(1-q_{s,x-1}).(1-q_{x-1}{}^{A})$
for $x = 1, 2, 3, ..., 13$ (and thus $r = 1$)

c) $F_{s,x,r,d}(t+1) = F_{s,x-1,r,d-1}(t).(1-q_{s,x-1}).(1-q_{d-1}{}^{A})$
for $x = 15, 16, 17, ..., 59$ (and thus $r = 2, 3, 4, 5$) and $x = 61, 62, 63, ..., 90$ (and thus $r = 6$)

d) $F_{s,x,r,d}(t+1) = \sum_{j=2}^{5} F_{s,x-1,j,d-1}(t).(1-q_{s,x-1}).(1-q_{d-1}{}^{A})$
for $x = 60$ (and thus $r = 6$) *(4.15)*

Note that all HIV-positive individuals aged x last birthday (where $x < 14$) are assumed to have been infected at birth, and their curtate duration since infection is thus equal to x.

4.4.4 Non-AIDS Deaths

Non-AIDS deaths can occur amongst HIV-negative and HIV-positive individuals.

The number of non-AIDS deaths occurring during year t amongst individuals with gender s, aged x last birthday and in risk group r at time t (denoted $D_{s,x,r}(t)^{N}$) is calculated as follows:

a) $D_{s,x,r}(t)^{N} = (C_{s,x,r}(t) + F_{s,x,r,x}(t)).q_{s,x}$
for $x = 0, 1, 2, ..., 13$ (and thus $r = 1$)

b) $D_{s,x,r}(t)^{N} = (C_{s,x,r}(t) + \sum_{d=0}^{19} F_{s,x,r,d}(t)).q_{s,x}$
for $x = 14, 15, 16, ..., 59$ (and thus $r = 2, 3, 4, 5$) and $x = 60, 61, 62, ..., 90$ (and thus $r = 6$) *(4.16)*

4.4.5 AIDS-Related Deaths

The number of AIDS-related deaths occurring during year t amongst individuals with gender s, aged x last birthday, in risk group r and with curtate duration since infection d at time t (denoted $D_{s,x,r,d}(t)^{A}$) is calculated as follows:

a) $D_{s,x,r,d}(t)^{A} = F_{s,x,r,x}(t).(1-q_{s,x}).\theta$
for $x = 0, 1, 2, ..., 13$ (and thus $r = 1$)

b) $D_{s,x,r,d}(t)^{A} = F_{s,x,r,d}(t).(1-q_{s,x}).q_{d}{}^{A}$
for $x = 14, 15, 16, ..., 59$ (and thus $r = 2, 3, 4, 5$) and $x = 60, 61, 62, ..., 90$

(and thus $r = 6$) *(4.17)*

Equation 4.17 effectively assumes that AIDS-related deaths occur after normal deaths have occurred. The error is slight.

4.5 Summarised Results

Figure 4.2 compares the HIV prevalence rates amongst antenatal clinic attendees and the HIV prevalence of pregnant women as projected by the model. The model was calibrated to the 1990 to 1995 prevalence rates, and appears to represent the 1996 prevalence rate satisfactorily.

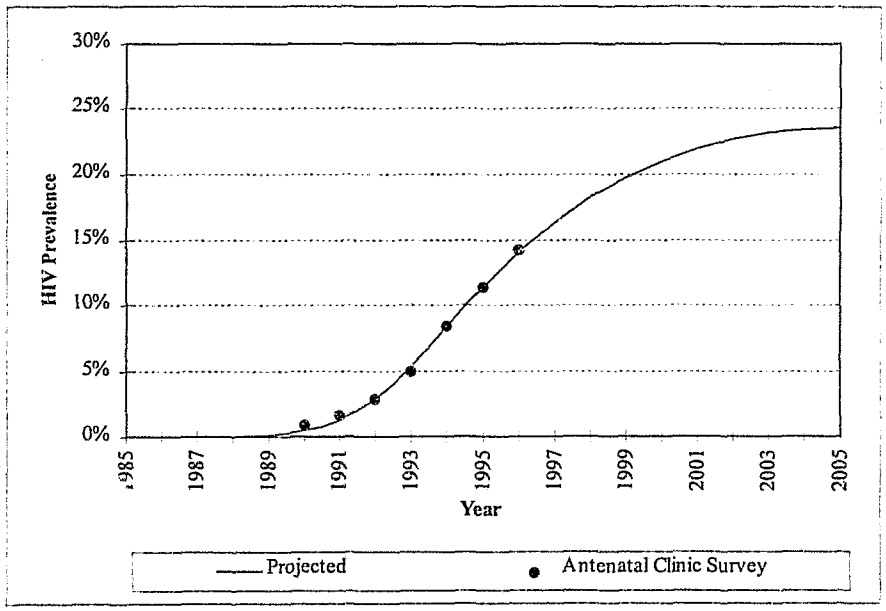


Figure 4.2 Comparison of the antenatal clinic data published by the Department of Health, and the HIV prevalence of pregnant women as projected by the ASSA model

Summarised projections under Scenario 500 of the ASSA model are shown in Tables 4.?

The ...iat approximately 2,5 million black South Africans were HIV-in... This is consistent with statistics released by the Department of Health, ...u the results of the seventh annual HIV survey conducted amongst antenatal clinic attendees (*Business Day*, 1997 and *The Star*, 1997). Just

under 4 million are projected to be infected by the turn of the century, reaching the 4,4 million mark by 2005. The number of females infected is consistently higher than the number of infected males.

The number of AIDS-related deaths is projected to have overtaken the number of normal (non-AIDS) deaths by the year 2005. More than 2,5 million AIDS-related deaths will have occurred by that time. Figure 4.3 illustrates the projected normal and AIDS-related deaths.

An extra 1,3% of the population is expected to die as a result of AIDS in the year 2005. A comparison of the mortality rates with and without the effect of AIDS is shown in Figures 4.4 and 4.5. Note that the presence of AIDS increases the mortality rates of children and the economically active population. The increase in mortality due to AIDS is markedly greater for females than for males.

Table 4.2 Projected statistics indicating the course of the AIDS epidemic in South Africa

	1991	1994	1996	2000	2005
HIV-positive males	97 489	643 982	1 084 878	1 752 564	2 008 170
HIV-positive females	133 055	851 612	1 415 684	2 180 991	2 393 226
Total HIV-positive	230 544	1 495 594	2 500 562	3 933 555	4 401 396
Deaths					
Normal deaths	323 609	344 328	358 548	386 264	411 295
AIDS-related deaths	3 086	24 999	60 897	224 128	451 655
Cumulative AIDS-related deaths	4 810	51 507	152 434	762 070	2 647 022
AIDS-related deaths					
Males aged 20-55	1 243	9 096	21 200	82 044	171 136
Females aged 20-55	1 038	8 146	21 501	99 686	214 864
Total aged 20-55	2 281	17 243	42 701	181 730	386 000
Crude extra mortality rate for total black population	0,0001	0,0008	0,0020	0,0067	0,0129

Table 4.3 Projected HIV prevalence rates amongst various subgroups of the population aged 15-45

Subgroup	1991	1994	1996	2000	2005
Female PRO	26,07%	82,82%	91,79%	97,90%	99,06%
Male PRO	19,18%	73,16%	81,07%	81,01%	76,08%
Female STD	3,44%	29,97%	51,14%	73,03%	82,48%
Male STD	2,40%	20,17%	35,56%	54,13%	59,98%
Female RSK	0,07%	1,13%	3,52%	10,94%	19,44%
Male RSK	0,05%	0,80%	2,45%	6,88%	11,09%
Female population	2,05%	12,21%	19,34%	27,41%	29,19%
Male population	1,47%	8,93%	14,16%	20,47%	21,65%

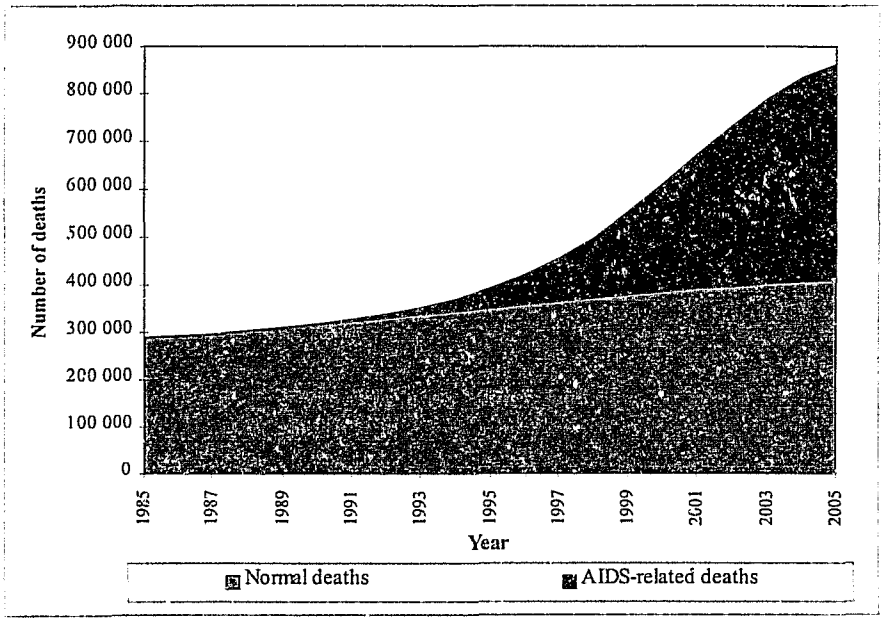


Figure 4.3 Projected normal and AIDS-related deaths

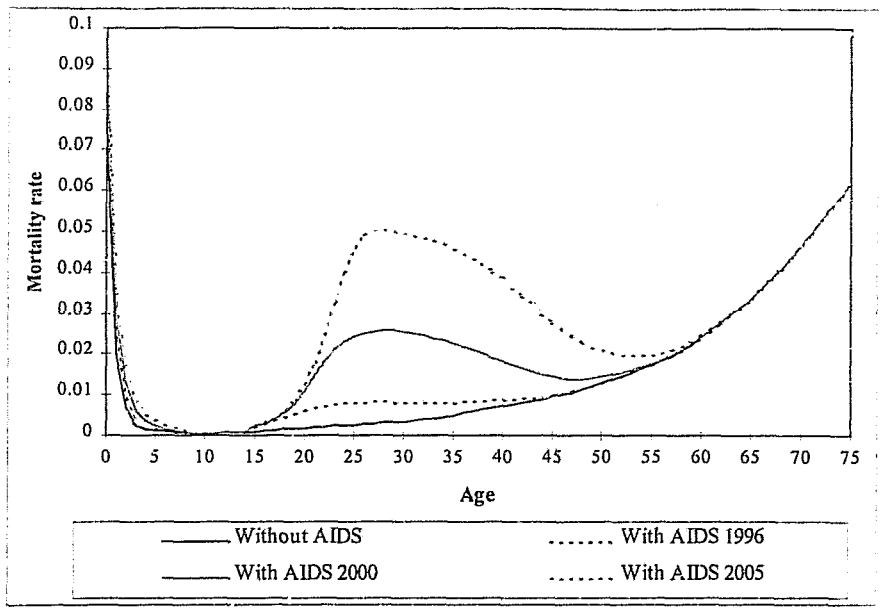


Figure 4.4 Projected female mortality curves, with and without the effect of AIDS

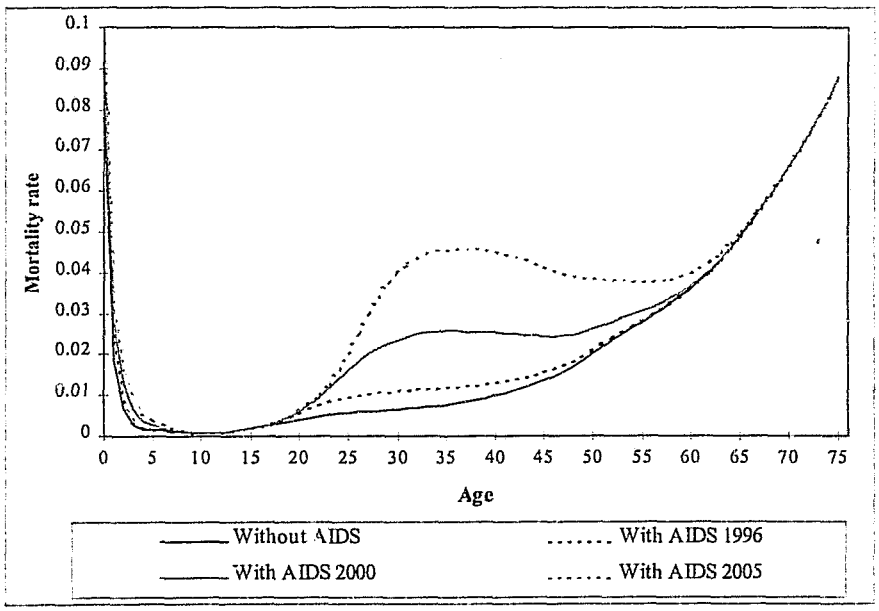


Figure 4.5 Projected male mortality curves, with and without the effect of AIDS

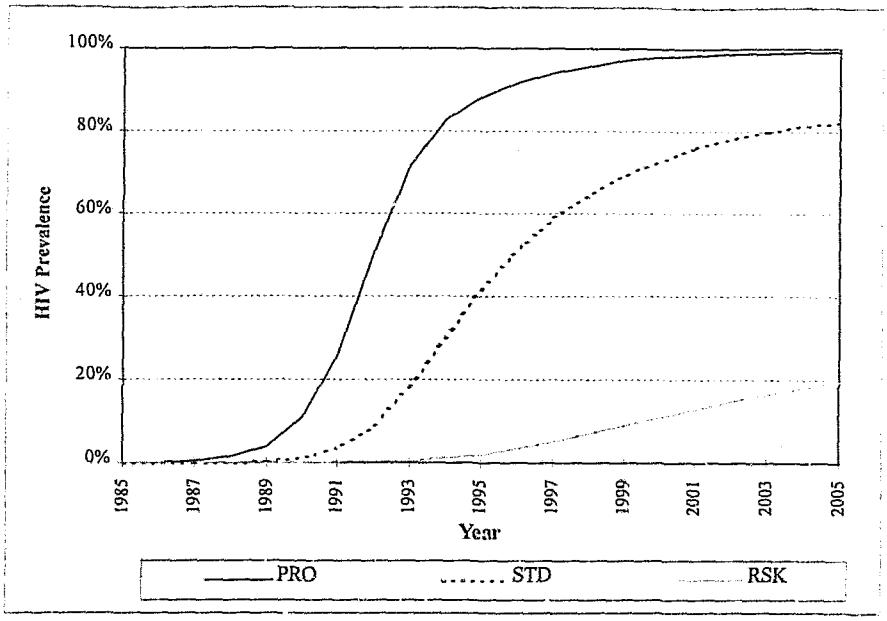


Figure 4.6 Projected HIV prevalence rates amongst females aged 15-45 in the PRO, STD and RSK risk groups

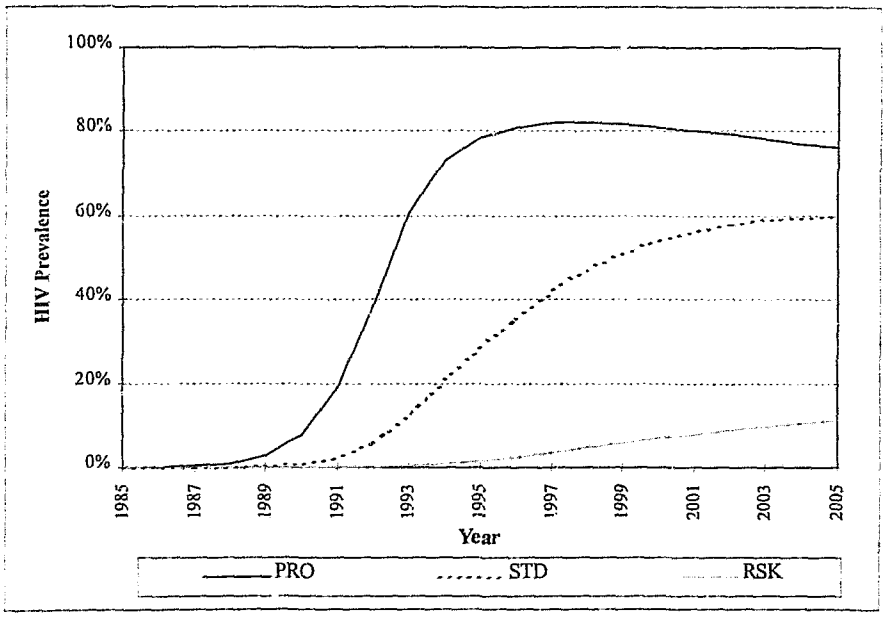


Figure 4.7 Projected HIV prevalence rates amongst males aged 15-45 in the PRO, STD and RSK risk groups

The prevalence rates shown in Table 4.3 indicate that a greater proportion of females than males are expected to be infected in each year, and within each subgroup. Almost all females within the PRO group are projected to be HIV-infected by the year 2005. The prevalence of the male PRO group is projected to decline after 1998. If the projection is extended beyond the year 2005, the prevalence gradually increases again, reaching 77% in 2025. This feature is a result of the different sexual activity indices for males and females. Approximately 27% of females and 20% of males between the ages of 15 and 45 are expected to be infected by the turn of the century. These prevalence rates are projected to increase to 29% and 22% by the year 2005. Figures 4.6 and 4.7 illustrate the prevalence rates for the PRO, STD and RSK risk groups.

4.6 Comparison with the Doyle Model

The ASSA model was developed as a basis for the recommendation of actuarial guidelines. The Doyle model had previously been used for this purpose, but was of limited use due to its proprietary nature. It is t' 's of interest to compare the results.

The structure of the ASSA model is similar to that of the Doyle model. In particular, similar assumptions are made regarding risk groups, and infected immigrants are assumed to start the epidemic in both models. The Doyle model differs from the ASSA model in a number of respects:

- It projects the spread of the epidemic amongst the total South African population, whereas the ASSA model only models the impact on the black population.
- It is calibrated against extensive data from Sub-Saharan Africa, whereas the ASSA model is primarily calibrated against the local antenatal clinic data.
- It uses assumptions regarding the probability of infection per contact and the number of new partnerships per year, rather than a "sexual activity weighted HIV prevalence".
- It differentiates between different stages of the disease, in particular HIV-infected individuals and people with AIDS. AIDS mortality is assumed to equal 50% per year for both adults and infants, whereas the ASSA model calculates AIDS mortality rates from a table of survival probabilities indexed by duration since infection.

Scenario 61 of the Doyle model projects that the number of HIV-infected individuals will increase from approximately 1 million in 1995 to 3,7 million in 2000, and then to 4,8 million in 2005. The ASSA model estimates a more gradual increase from approximately 2 million in 1995 to 3,9 million in 2000 and 4,4 million in 2005. The ASSA model seems to correspond to the estimated 2,5 million infected South Africans in 1996 as released by the Department of Health (*Business Day*, 1997 and *The Star*, 199'/).

The number of AIDS-related deaths projected by the ASSA model falls between those estimated by Scenarios 61 and 60 of the Doyle model. In particular, the ASSA model predicts that approximately 450 000 AIDS-related deaths will occur in 2005, compared to the 430 000 and 525 000 projected under Scenarios 61 and 60 respectively.

The HIV prevalence rates amongst adults projected by Doyle's Scenario 60, increases from approximately 5% in 1995 to 24% in 1995. The ASSA model predicts a higher prevalence rate in the early 1990s (14% in 1995) to a similar level at the later stages of the epidemic (25% in 2005).

4.7 Summary

The ASSA AIDS model was produced so as to be of service to the actuarial profession, as well as to the public at large. The workings are publicly available, although one has to study the spreadsheet and Visual Basic code in order to deduce how the projections are actually made.

This chapter presents a detailed description of the model. In particular, it discusses the background to the development of the model, gives an overview of the model structure, discusses the various inputs, outlines the projections in detail and provides summarised results.

These results can be compared to the results of the calibrated model shown in Chapter 7.

5

Principal Estimation of AIDS-Related Deaths

The estimation of AIDS-related deaths is fraught with difficulties (Botha and Bradshaw, 1985; Bah and Kleinschmidt, 1997). Firstly, a significant number of deaths (especially black deaths) are not reported to the authorities. Secondly, recorded deaths may be attributed to an incorrect, or vague cause. This point is particularly relevant to AIDS-related deaths, as doctors and/or relatives may deliberately disguise the true cause of death to avoid possible stigma, or compromising insurance claims.

This chapter describes the estimation of AIDS-related deaths. Various techniques were required to overcome the inadequacies of the data.

5.1 Overview of Method

Mortality normally changes gradually over time. However, the AIDS epidemic is expected to alter the shape of the conventional mortality curve relatively rapidly. This change in the shape of the mortality curve was used to estimate AIDS-related deaths.

The principal estimation of AIDS-related deaths was conducted exclusively for the black population for two reasons:

- The AIDS epidemic is expected to be rife amongst the black population, and relatively insignificant amongst the other three population groups.
- The estimates are to be used to calibrate the ASSA model, which projects the spread of the AIDS epidemic amongst the black population.

It was also necessary to select a period of investigation. This period had to be as recent as possible since AIDS-related deaths have only recently started emerging. Data on reported deaths within the South African population are available for years up to and including 1994. The investigation period was thus defined as a four-year period, from 1991 to 1994.

The following steps were taken in order to compute age and sex specific mortality rates for the black population:

- Extraction of black reported deaths from the total reported deaths. This was necessary since death information was not available by population group. This step is discussed in section 5.3.1.
- Adjustment of black reported deaths to allow for underreporting. Two methods are used to estimate the completeness of reporting of black deaths: the Preston and Coale method (described in section 5.3.2) and the Bennett and Horiuchi method (described in section 5.3.3).
- Calculation of actual mortality rates by dividing the estimated actual deaths by a suitable exposed. This is discussed in sections 5.4 and 5.5.
- Choice of a suitable "expected" mortality curve, and adjustment of this curve to eliminate changes in the level of mortality. Section 5.6 deals with this issue.
- Estimation of AIDS-related deaths by comparing actual and expected mortality rates, and attributing excess deaths to AIDS. This is discussed in section 5.7.

Each of the above steps involves a number of problems.

The method as a whole relies on the assumption that changes in the shape of the mortality curve are a result of AIDS. In reality, however, a number of other factors may influence the shape of the mortality curve. Examples of such factors include violence (which is expected to have been particularly significant amongst the black population in the run-up to the 1994 election), mortality improvements over certain age ranges due to an improved standard of living, and increased mortality due to motor vehicle accidents. Individual causes will be analysed in greater detail in chapter 6 to provide some justification for this important assumption.

5.2 Notation

The following notation will be used.

- Let $i = 1, 2, 3, 4$ denote asians, coloureds, whites, and blacks respectively.
- Let $s = 0, 1$ denote males and females respectively.
- Let $t = 1, 2, 3, 4, 5$ denote the start of calendar year $1990+t$.
- Let ${}_{y-x}N_{t,i,s,x}$ denote the number of individuals in population group i, with gender s and aged between x and y on 30th June of year t.
- Let ${}_{y-x}D_{t,i,s,x}$ denote the number of deaths which occurred during year t amongst individuals in population group i, with gender s and aged between x and y.
- Let ${}_{y-x}D_{t,i,s,x}^{R}$ denote the number of reported deaths which occurred during year t amongst individuals in population group i, with gender s and aged between x and y.
- Let $B_{t,i,s}$ denote the number of babies born into population group i and with gender s during the year from time t-1/2 to time t+1/2.

- Let ${}_{y-x}AIDS_{t,i,s,x}$ denote the number of AIDS-related deaths which occurred in year t amongst individuals in population group i, with gender s and aged between x and y.
- Let ${}_{y-x}f_{t,i,x}$ denote the age-specific fertility rate for women in population group i, aged between x and y which is applicable for the year from time t-1/2 to time t+1/2.
- Let ${}_{y-x}q_{t,i,s,x}{}^{A}$ denote the actual average probability for year t that an individual in population group i, with gender s and aged between x and y will die within one year.
- Let ${}_{y-x}q_{t,i,s,x}{}^{E}$ denote the expected average probability for year t that an individual in population group i, with gender s and aged between x and y will die within one year.
- Let ${}_{y-x}q_{t,i,s,x}{}^{AIDS}$ denote the estimated probability for year t that an individual in population group i, with gender s and aged between x and y will die as a result of AIDS within one year.
- Let ${}_{y-x}q_{t,s,x}{}^{NonAIDS}$ denote the estimated probability for year t that an individual in population group i, with gender s and aged between x and y will die as a result of non-AIDS causes within one year.
- Let $l_{i,s,x}{}^{S}$ denote the number of individuals in population group i and with gender s who attain age x according to a standard table. Note that $l_{i,s,0}{}^{S}$ is arbitrarily defined.
- Let ${}_{y-x}l_{i,s,x}{}^{S} = \sum_{a=x}^{y-1} l_{i,s,a}$.
- Let $q_{i,s,x}{}^{S}$ denote the probability that an individual in population group i, with gender s and age x will die within one year, according to a standard table.
- Let ${}_{y-x}q_{i,s,x}{}^{S} = \frac{1}{{}_{y-x}l_{i,s,x}{}^{S}} \sum_{a=x}^{y-1} q_{i,s,a}{}^{S} . l_{i,s,a}{}^{S}$ denote the average probability that an individual in population group i, with gender s and aged between x and y will die within one year, according to a standard table.

A single bar will be used to indicate that a variable refers to all population groups combined, and the subscript referring to population group (i) will then be dropped. For example,

$${}_{y-x}\overline{N}_{t,s,x} = \sum_{i=1}^{4} {}_{y-x}N_{t,i,s,x}$$

denotes the number of individuals in all population groups with gender s and aged between x and y on 30th June of year t.

A double bar will be used to indicate that a variable refers to all population groups and both genders combined. The subscripts referring to population group (i) and gender (s) will then be dropped. For example,

$$_{y-x}\overline{\overline{N}}_{t,x} = \sum_{i=1}^{4}\sum_{s=0}^{1} {}_{y-x}N_{t,i,s}$$

denotes the total number of individuals aged between x and y on 30th June of year t.

5.3 Estimation of the Actual Number of Black Deaths Which Occurred in 1991, 1992, 1993 and 1994

The Central Statistical Service (CSS) publishes an annual report which provides demographic and other information on deaths reported during the year. Information by population group has not been available since 1991.

The 1991, 1992, 1993 and 1994 death reports (CSS, 1991 a; CSS, 1992; CSS, 1993 b and CSS, unpublished) provide information on deaths reported during each of these years. Deaths arising in the former TBVC states were excluded until 1994. It was thus implicitly assumed that deaths arising in the former TBVC states in 1992 and 1993 had similar age and gender distributions to deaths arising in the provinces during these years. This may not be true, and the assumption may introduce inaccuracy into the results.

Reporting delays imply that deaths reported during year t may have occurred within a prior year.

Limited information is available on deaths by year of death. The information which is available (Bah: Personal communication) is shown in Table 5.1.

Table 5.1 Deaths grouped by year of death and year of registration

Year of death	Number of years until death is registered			
	0	1	2	3 or more
1990				x_3
1991			x_1	220
1992		10 350	484	
1993	187 764	2 895	x_2	x_4
1994	209 680			

Source: Central Statistical Service

Using Table 5.1 and data on deaths grouped by year of reporting, the following observations can be made:

- 93,29% of deaths reported in 1993 actually occurred during that year, whereas 98,31% of deaths reported in 1994 actually occurred during that year.
- 5,14% of deaths reported in 1993 actually occurred in 1992, whereas 1,36% of deaths reported in 1994 actually occurred in 1993.
- 1,57% of deaths reported in 1993 actually occurred some time before 1992, whereas 0,33% of deaths reported in 1994 actually occurred some time before 1993.
- $x_1 + x_3 = 3159$.

Reporting delays have improved between 1993 and 1994. The number deaths reported in 1993 includes the figure 10 350 instead of 2 895, as well as x_1 instead of x_2 and x_3 instead of x_4, as shown in Table 5.1. If reporting delays are assumed to have improved between each year t and t+1, then x_1 is greater than x_2, and x_3 greater than x_4. This would imply that the number of deaths reported during 1993 is an overestimate of the number of reported deaths which occurred during 1993.

So, if it is assumed that reporting delays have improved each year, then the number of deaths reported during year t overstates the number of reported deaths which occurred in year t. This error decreases with t.

It would be preferable to use death data grouped by year of death rather than year of reporting. However, in the absence of such data, reporting delays will be ignored. It is thus assumed that all deaths reported in year t occurred in year t. The distortion caused by this approximation must be borne in mind when interpreting the results.

5.3.1 Extraction of Black Reported Deaths from Total Reported Deaths

The actual pattern of asian, coloured and white deaths was assumed to equal an expected pattern based on standard tables of mortality rates. These rates were used, together with estimates of the asian, coloured and white populations, to extract the black reported deaths from the aggregate reported deaths.

Standard Tables for the Asian, Coloured and White Population Groups

The CSS has produced a set of life tables for the asian, coloured and white population groups based on the 1985 census results and death statistics for the years 1984-1986 (CSS, 1987). These tables, called the SA 84-86 tables, refer to a period approximately ten years ago, and may thus be out of date. However, in the absence of more recent population life tables, they have been chosen as the

standard tables for the asian, coloured and white population groups. The statistics denoted ${}_{i,s,x}l^S$ and ${}_{i,s,x}q^S$ (for $i = 1, 2, 3$) thus refer to these standard tables.

Extrapolation of the Standard Tables. The mortality rates, ${}_{i,s,x}q^S$, are only defined for ages less than 90 in the SA 84-86 tables. The rates were thus first extrapolated to an extreme age ω (where $\omega = 115$) by fitting the following model to ${}_{i,s,88}q^S$ and ${}_{i,s,89}q^S$ in each standard table:

$$\frac{{}_{i,s,x}q^S}{1-{}_{i,s,x}q^S} = B_{i,s}.c_{i,s}{}^x \qquad (5.1)$$

where $B_{i,s}$ and $c_{i,s}$ are constants

This model is a special case of models proposed by Perks (1932), and Heligman and Pollard (1980).

The extrapolated mortality curves are shown in Figures 5.1 and 5.2.

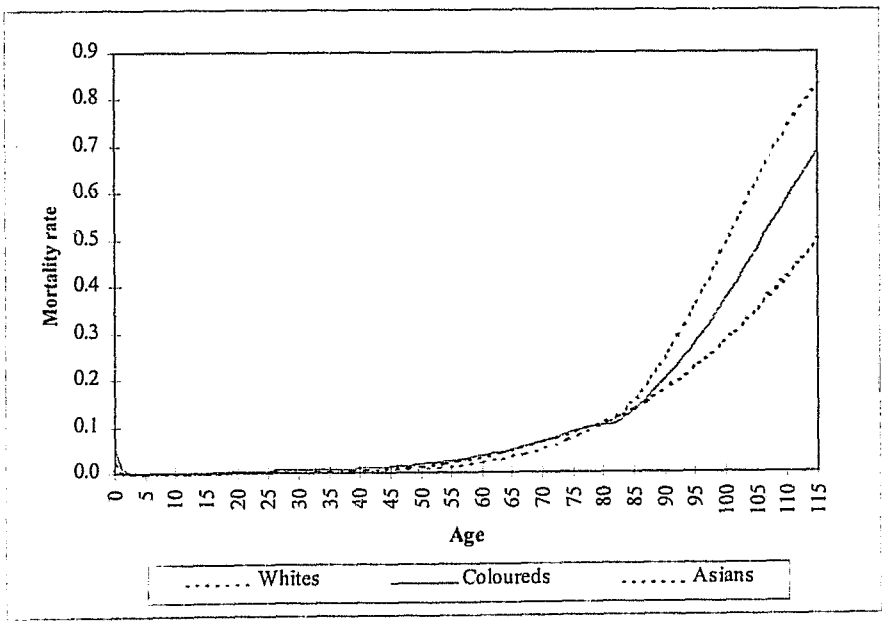


Figure 5.1 Extrapolated standard male mortality curves

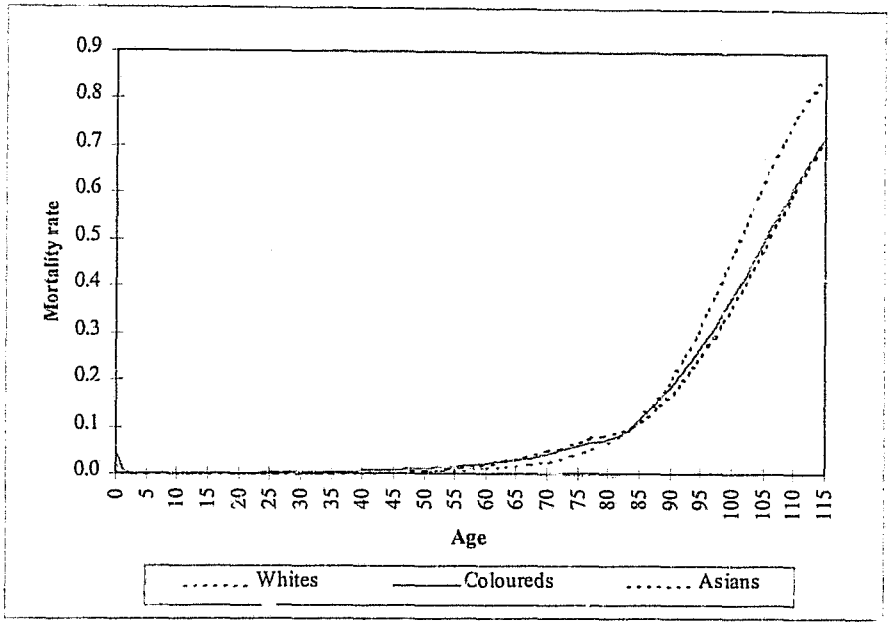


Figure 5.2 Extrapolated standard female mortality curves

Mortality Crossovers. Mortality is generally expected to be heaviest for those sectors of the population which are disadvantaged with respect to a variety of social, economic and public health variables. It is therefore expected that the white population in South Africa would experience lighter mortality than asians and coloureds.

The mortality curves from the standard tables shown in Figures 5.1 and 5.2 follow this expected pattern for all ages up to age y (where $y \approx 83$). From age y onwards, however, white mortality appears to be heaviest. This phenomenon is called a mortality crossover. Such a crossover may exist in reality, or may be an artificial phenomenon resulting from data inaccuracies.

Mortality crossovers have been documented in the past. A black-white mortality crossover occurs at age 75 in the US 1969-1971 life tables (NCHS, 1975). Crossovers have been observed in international comparisons of the mortality experience of selected human populations (Nam, Weatherby and Ockay, 1978). Cause-specific mortality crossovers have also been observed (Manton, Poss and Wing, 1979).

A number of observed mortality crossovers were discussed in a paper by Manton and Stallard (1984:246-258). Possible explanations for this phenomenon were presented:

- Those individuals from the disadvantaged sector of the population who survive to advanced ages are constitutionally endowed for longer survival. The "weaker" individuals die at relatively younger ages, and only the "hardy" individuals remain. These individuals experience very light mortality, and hence, a mortality crossover occurs.

- Individuals from the disadvantaged sector of the population have a different pattern of ageing to the rest of the population. These individuals may age more rapidly, and thus experience heavier mortality, at the middle ages. The rest of the population may age more rapidly at advanced ages, and thus experience heavier mortality at these ages. A mortality crossover would thus be observed.

- Individuals from the disadvantaged sector of the population experience heavier mortality at the middle ages, since these are the ages at which they are exposed to the most hazardous environmental conditions. A mortality crossover would occur if this environmental stress lessened at the older ages, to the extent that these individuals experienced less stress at the older ages than the rest of the population.

- The disadvantaged sector of the population tends to be overenumerated at older ages. Deaths occurring in this sector tend to be underreported at the older ages. These two sources of data error result in an artificial mortality crossover.

Adjustment of Standard Mortality Tables. The final point is a particular problem in South Africa where the state old-age pension paid to those over 65 represents a significant proportion of family income, and there is thus great financial incentive to overstate age and not report deaths. It is thus a likely explanation for the observed mortality crossover. The standard tables were thus adjusted to eliminate this phenomenon.

The standard table chosen for the white population was believed to be the most reliable. Mortality rates at the older ages for the asian and coloured population groups were therefore adjusted in line with this table.

Instead of a mortality crossover, a convergence in mortality rates was assumed. This assumption gives some credit to the empirical evidence whilst eliminating the probable distortions produced by unreliable data. In particular, mortality rates for the asian and coloured population groups were assumed to converge towards white mortality rates from age 76. All population groups were assumed to experience the same mortality from age 100.

The methods used to adjust the standard tables were designed so as to meet two criteria:

- The resulting mortality curves should be as smooth as possible.

- Age-specific mortality rates should generally be ranked from heaviest to lightest as follows: $q_{2,s,x}{}^S$, $q_{1,s,x}{}^S$, $q_{3,s,x}{}^S$.

The *white* standard table was extrapolated as described above, but by fitting the model to $q_{3,s,74}{}^S$ and $q_{3,s,75}{}^S$ for males, and $q_{3,s,72}{}^S$ and $q_{3,s,73}{}^S$ for females. These mortality rates were chosen as they yielded smooth mortality curves.

The *asian* standard table was adjusted as follows for males:

$$\begin{aligned} &q_{1,0,x}{}^S \text{ unchanged} && \text{for } x < 76 \\ &q_{1,0,x}{}^S = q_{3,0,x}{}^S + m_{1,0}(100\text{-}x) && \text{for } x = 76, 77, ..., 100 \\ &q_{1,0,x}{}^S = q_{3,0,x}{}^S && \text{for } x > 100 \end{aligned} \qquad (5.2)$$

where $m_{1,0} = (q_{1,0,75}{}^S - q_{3,0,75}{}^S)/25$

The *asian* standard table was adjusted as follows for females:

$$\begin{aligned} &q_{1,1,x}{}^S \text{ unchanged} && \text{for } x < 61 \\ &q_{1,1,x}{}^S / (1 - q_{1,1,x}{}^S) = B_{1,1} c_{1,1}{}^x && \text{for } x = 61, 62, ..., 75 \\ &q_{1,1,x}{}^S = q_{3,1,x}{}^S + m_{1,1}(100\text{-}x) && \text{for } x = 76, 77, ..., 100 \\ &q_{1,1,x}{}^S = q_{3,1,x}{}^S && \text{for } x > 100 \end{aligned} \qquad (5.3)$$

where $B_{1,1}$ and $c_{1,1}$ are constants determined by fitting equation 5.1 to $q_{1,1,59}{}^S$ and $q_{1,1,60}{}^S$,
and $m_{1,1} = (q_{1,1,75}{}^S - q_{3,1,75}{}^S)/25$

The *coloured* standard table was adjusted as follows for males:

$$\begin{aligned} &q_{2,0,x}{}^S \text{ unchanged} && \text{for } x < 76 \\ &q_{2,0,x}{}^S = q_{3,0,x}{}^S + m_{2,0}(100\text{-}x) && \text{for } x = 76, 77, ..., 100 \\ &q_{2,0,x}{}^S = q_{3,0,x}{}^S && \text{for } x > 100 \end{aligned} \qquad (5.4)$$

where $m_{2,0} = (q_{2,0,75}{}^S - q_{3,0,75}{}^S)/25$

The *coloured* standard table was adjusted as follows for females:

$$\begin{aligned} &q_{2,1,x}{}^S \text{ unchanged} && \text{for } x < 76 \\ &q_{2,1,x}{}^S = q_{3,1,x}{}^S + m_{2,1}(100\text{-}x) && \text{for } x = 76, 77, ..., 100 \\ &q_{2,1,x}{}^S = q_{3,1,x}{}^S && \text{for } x > 100 \end{aligned} \qquad (5.5)$$

where $m_{2,1} = ({}_{75}q_{2,1}^{S} - {}_{75}q_{3,1}^{S})/25$

The adjusted mortality curves are shown in Figures 5.3 and 5.4. The rates at the older ages are reasonably consistent with the 1991 census enumeration. Differences may be attributable to enumeration error and (limited) migration.

Estimates of the Asian, Coloured and White Populations

Estimates of the size and demographic structure of the asian, coloured and white population groups were required for each of the years 1991 to 1994. The published 1991 census results (CSS, 1991 c), denoted ${}_{y\text{-}x}N_{1,i,s,x}$, were assumed to reflect the size of the asian, coloured and white population groups at the midpoint of 1991, although the census was actually conducted on 7th March 1991. The census results group the population into five-year age intervals up to age 90, and a final open interval from age 90 onwards.

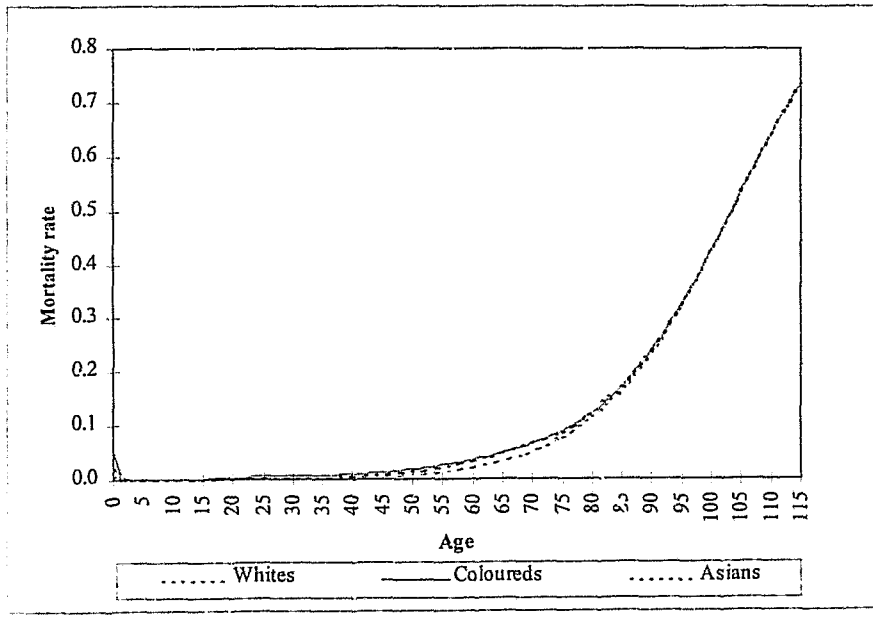


Figure 5.3 Adjusted standard male mortality curves

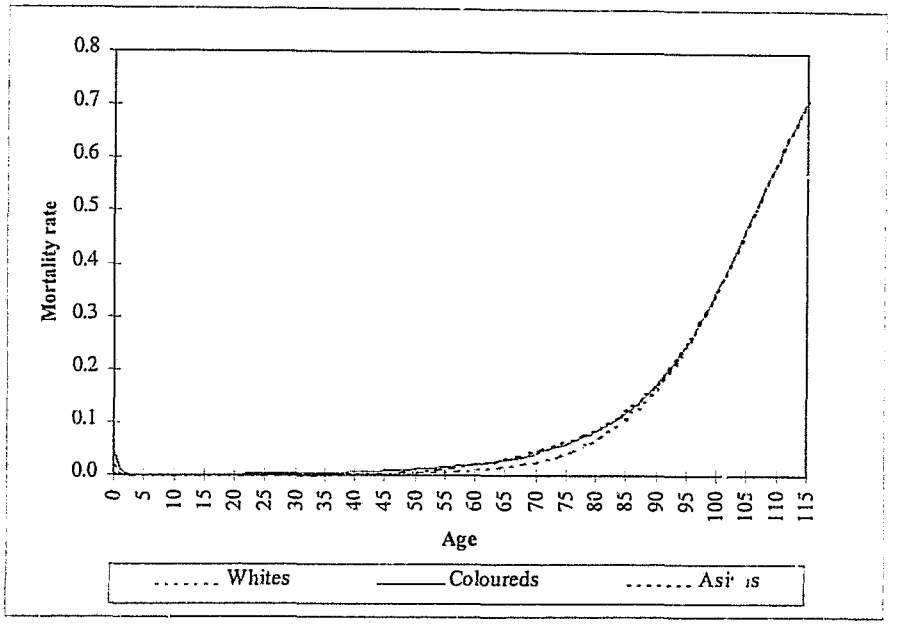


Figure 5.4 Adjusted standard female mortality curves

The estimated population at the midpoint of year t+1 (t = 1, 2, 3) was calculated as follows:

$$ {}_5N_{t+1,i,s,x} = \left(\frac{{}_4l^{S}_{i,s,x+1}}{{}_5l^{S}_{i,s,x}}\right)\cdot {}_5N_{t,i,s,x} + B_{t+1,i,s} \qquad \text{for } x = 0 $$

$$ {}_5N_{t+1,i,s,x} = \left(\frac{{}_4l^{S}_{i,s,x+1}}{{}_5l^{S}_{i,s,x}}\right)\cdot {}_5N_{t,i,s,x} + \left(\frac{l^{S}_{i,s,x}}{{}_5l^{S}_{i,s,x-5}}\right)\cdot {}_5N_{t,i,s,x-5} \qquad \text{for } x = 5, 10, \ldots, 85 $$

$$ {}_{\omega-x}N_{t+1,i,s,x} = \left(\frac{{}_{\omega-x-1}l^{S}_{i,s,x+1}}{{}_{\omega-x}l^{S}_{i,s,x}}\right)\cdot {}_{\omega-x}N_{t,i,s,x} + \left(\frac{l^{S}_{i,s,x}}{{}_5l^{S}_{i,s,x-5}}\right)\cdot {}_5N_{t,i,s,x-5} \qquad \text{for } x = 90 \quad \textit{(5.6)} $$

where ${}_{y-x}l_{i,s,x}{}^{S}$ and ${}_{y-x}q_{i,s,x}{}^{S}$ refer to the standard mortality tables discussed in the previous section.

Studies have shown that 52% of babies born are male (Newell, 1988:27-30). However, the ratio tends to vary by population group, and is estimated to be lower for blacks (Udjo, 1997). The ratio for the asian, coloured and white population groups was assumed to be close to 52%.

$B_{t,i,s}$ was thus calculated as shown in equation 5.7:

$$B_{t,i,s} = 0{,}52\sum_{x=15}^{21} {}_5N_{t,i,1,5(x-15)+15}\cdot {}_5f_{t,i,5(x-15)+15} \quad \text{for males}$$

$$B_{t,i,s} = 0{,}48\sum_{x=15}^{21} {}_5N_{t,i,1,5(x-15)+15}\cdot {}_5f_{t,i,5(x-15)+15} \quad \text{for females} \qquad (5.7)$$

The values of $f_{t,i,x}$ were set equal to the fertility rates used to adjust the 1991 census (CSS, 1991 b). Females were assumed to fall pregnant between the ages of 15 and 50.

Estimated Number of Black Reported Deaths

It was assumed that $100p\%$ of the asian, coloured and white deaths were reported, and that the remaining deaths were black.

The number of black reported deaths was calculated as shown in equation 5.8:

$${}_{y-x}D_{t,4,s,x}{}^{R} = {}_{y-x}\overline{D}_{t,s,x}{}^{R} - p\sum_{i=1}^{3} {}_{y-x}N_{t,i,s,x}\cdot {}_{y-x}q_{i,s,x}{}^{S} \qquad (5.8)$$

where values of ${}_{y-x}\overline{D}_{t,s,x}{}^{R}$ were obtained from the death reports.

Three scenarios were constructed:

- **Scenario 1:** $p = 100\%$
- **Scenario 2:** $p = 90\%$
- **Scenario 3:** $p = 80\%$

The estimated number of reported black deaths under the three scenarios appear in Table 5.2. A breakdown by age and gender under Scenario 2 is shown in Appendix A.

Table 5.2 *Estimated number of reported black deaths*

Year	Scenario 1		Scenario 2		Scenario 3	
	Males	Females	Males	Females	Males	Females
1992	59 201	38 243	63 561	41 922	67 921	45 602
1993	72 952	45 970	77 417	49 740	81 882	53 510
1994	77 808	51 185	82 377	55 045	86 945	58 905

Possible Distortions

The estimates shown in Table 5.2 may be distorted for the following reasons:

- The actual pattern of asian, coloured and white deaths is unlikely to equal the expected pattern based on standard tables of mortality rates. Any difference between the actual and expected mortality curves would affect the total number of asian, coloured and white deaths within each age group, and is likely to distort the number and proportion of black deaths within each age group.

- The completeness of reporting of asian, coloured and white deaths is unlikely to be uniform across all age groups. In particular, deaths occurring amongst adults are more likely to be reported than deaths occurring at the very old and very young ages. The assumption of uniform completeness of reporting may thus underestimate the number of reported black deaths at the extreme ages.

5.3.2 The Preston and Coale Method

Assumptions

The so-called Preston and Coale method (Preston *et al*, 1980 and UN, 1993) is used to estimate the completeness of reporting of deaths. It relies upon two assumptions:

a) The population under observation is stable.
b) The completeness of reporting is uniform across all ages.

A stable population is one which grows at the same annual rate at all ages. Births and deaths grow at this same rate. Such a population has experienced constant fertility and mortality for a long period of time. Few populations are stable, and assumption (a) may distort the results.

Assumption (b) is unlikely to be true, since the reporting of deaths amongst very young and very old individuals tends to differ from that of the rest of the population. The method is thus conventionally used to estimate the completeness of reporting of the adult population only.

An Overview of the Mathematical Structure

Since all individuals will eventually die,

$$N_{t,i,s,x} = \sum_{a=x}^{\omega} D_{t+a-x,i,s,a} \qquad (5.9)$$

Now $D_{t+a-x,i,s,a}$ can be approximated as follows if the population is stable:

$$D_{t+a-x,i,s,a} = {D_{t,i,s,a}}^{R}.exp(r(a-x)) \qquad (5.10)$$

where r denotes the annual *force* of growth of the population under observation. Note that the corresponding annual *rate* of growth would equal $exp(r)-1$.

The Preston and Coale estimate of $N_{t,i,s,x}$ (denoted $\hat{N}_{t,i,s,x}$) can be calculated as shown in equation 5.11:

$$\hat{N}_{t,i,s,x} = \sum_{a=x}^{\omega} {D_{t,i,s,a}}^{R}.\exp(r(a-x)) \qquad (5.11)$$

Three sets of ratios may be calculated as an indication of the proportion of deaths which are reported:

$${c_{t,i,s,x}}^{1} = \hat{N}_{t,i,s,x} / N_{t,i,s,x} \qquad (5.12)$$

$${c_{t,i,s,x}}^{2} = {}_5\hat{N}_{t,i,s,x} / {}_5N_{t,i,s,x} \qquad (5.13)$$

$${c_{t,i,s,x}}^{3} = \sum_{a=x}^{(Y-5+4x)/5} {}_5\hat{N}_{t,i,s,5(a-x)+x} \Big/ \sum_{a=x}^{(Y-5+4x)/5} {}_5N_{t,i,s,5(a-x)+x} \text{ for some } Y \qquad (5.14)$$

Plots of each of these ratios against x should yield horizontal lines when assumptions (a) and (b) hold. r may be adjusted so as to yield the most horizontal lines.

The ratios ${c_{t,i,s,x}}^{1}$ tend to be more erratic than ${c_{t,i,s,x}}^{2}$, which in turn tend to be more erratic than ${c_{t,i,s,x}}^{3}$. The Preston and Coale estimate of the completeness of reporting of deaths occurring during year t amongst individuals in population group i and with gender s (denoted by $C_{t,i,s}$) is generally defined as the mean or median of the ratios ${c_{t,i,s,x}}^{3}$ over the adult ages (e.g. ages 20 to 60).

Approximations Made in Practice

Data on deaths are often grouped into five-year age bands (${}_5{D_{t,i,s,x}}^{R}$), with a final open interval (from some age X to a limiting age ω). Equation 5.11 therefore cannot be applied directly.

$\hat{N}_{t,i,s,x}$ may be calculated recursively assuming that reported deaths are linear within each quinquennial age band, as shown in equation 5.15:

$$\hat{N}_{t,i,s,x} = \hat{N}_{t,i,s,x+5}.\exp(5r)+{}_5D_{t,i,s,x}{}^{R}.\exp(2{,}5r) \qquad (5.15)$$

$\hat{N}_{t,i,s,X}$ is used as a starting point, where X is the lower bound of the open interval. It has been suggested (UN, 1993) that X should be set at a level which provides an adequate balance between the advantage of reducing the misreporting of age at the very old ages, and the disadvantage of approximating the weight of the deaths in this open interval.

The number of individuals aged X can be calculated using equation 5.16, where $z(X)$ is some length of time:

$$\hat{N}_{t,i,s,X} = (\sum_{a=X}^{\omega} D_{t,i,s,a}{}^{R})\exp(r.z(X)) \qquad (5.16)$$

Values of $z(X)$ were calculated for a number of model populations with different growth rates and different mortality experiences (UN, 1993). Least-squares regression was used to fit the following equation for each pattern (north, south, east and west) of the Coale-Demeny (1966) life tables:

$$z(X) = a(X) + b(X)r + c(X)\exp(\frac{\sum_{a=45}^{\omega} D_{t,i,s,a}{}^{R}}{\sum_{a=10}^{\omega} D_{t,i,s,a}{}^{R}}) \qquad (5.17)$$

where $a(X)$, $b(X)$ and $c(X)$ are constants. The west family of life tables was noted to be a satisfactory default.

As an alternative to equation 5.17, $z(X)$ can be calculated as follows:

$$z(X) = \frac{1}{r}.ln[\frac{\sum_{a=X}^{\omega} d_{i,s,a}{}^{S}.e^{r(a-X)}}{\sum_{a=X}^{\omega} d_{i,s,a}{}^{S}}] \qquad (5.18)$$

where $d_{i,s,x}{}^{S} = l_{i,s,x}{}^{S} - l_{i,s,x+1}{}^{S}$ from a standard life table. Equation 5.18 can be derived from equation 5.16.

Population data are often grouped into five-year age bands (${}_5N_{t,i,s,x}$). Corresponding estimates of the population can be obtained using the

approximation shown in equation 5.19:

$$_5\hat{N}_{t,i,s,x} = 2{,}5(\hat{N}_{t,i,s,x} + \hat{N}_{t,i,s,x+5}) \qquad (5.19)$$

Possible Distortions

The results of the Preston and Coale method may be distorted for a number of reasons:

- The completeness of reporting of black deaths may not be uniform even over the adult ages. The curves of equations 5.12, 5.13 and 5.14 would thus not be horizontal across the adult age range. Any attempt to make these curves horizontal may introduce inaccuracy and bias.

- The age of older individuals in the population may be misreported. It has been found that age is commonly overstated in the black population (Dorrington: Personal communication). This increases the size of the denominators in equations 5.12, 5.13 and 5.14 at the older ages, and thus reduces the ratios. Any age misreporting at ages below age X would distort the curve of equation 5.14 at all ages (with the distortion reducing as age decreases) due to the cumulative nature of the ratios. However, if X is chosen so as to restrict all age misstatement to the open interval and $C_{t,i,s}$ is calculated over an age range excluding ages greater than X (i.e. $Y = X$ in equation 5.14), then no distortion is expected (unless $z(X)$ is overstated).

- A decline in fertility alters the age distribution of the population and assumption (a) of the Preston and Coale method is thus violated. It has the effect of reducing the size of the young population, thus reducing the denominators in equations 5.12, 5.13 and 5.14 at these ages and increasing the ratios. However, if the fertility decline has not occurred more than 15 years prior to the investigation, and $C_{t,i,s}$ is derived by analysing ratios over age 15, then the results should be unbiased.

- An improvement in mortality, especially if not uniform across all ages, will cause a departure from the assumption of the stable age distribution of the population. This would distort the ratios in equations 5.12, 5.13 and 5.14.

- The chosen growth rate may be incorrect. A growth rate which is too low would result in an increasing sequence of $c_{t,i,s,x}$[3] ratios, whereas an overstated growth rate would lead to a decreasing sequence of ratios.

It is likely that a number of the abovementioned distorting factors operate simultaneously. For this reason, it may be difficult to distinguish between the various factors.

5.3.3 The Bennett and Horiuchi Method

This method was developed by Bennett and Horiuchi (1981) as an extension to the Preston and Coale method.

Whereas the latter relies upon the assumption of a constant population growth rate, the Bennett and Horiuchi method allows for population growth rates which vary with age and gender. Assumption (a) (section 5.3.2) of the Preston and Coale method is thus stated less restrictively as follows: The number of individuals in population group i, with gender s and aged between x and y increases annually according to a constant force of growth ${}_{y-x}r_{i,s,x}$. Note that the corresponding annual *rate* of growth would equal $exp({}_{y-x}r_{i,s,x})-1$.

Assumption (b) of the Preston and Coale method is still required.

Values of ${}_{y-x}r_{i,s,x}$ are usually obtained from censuses conducted in year t_1 and t_2, as shown in equation 5.20:

$$ {}_{y-x}r_{i,s,x} = \frac{1}{t_2 - t_1} ln[1 + (\frac{{}_{y-x}N_{t_2,i,s,x}}{{}_{y-x}N_{t_1,i,s,x}} - 1)] \qquad (5.20) $$

However, relative underenumeration or overenumeration in the second census would understate or overstate the age-specific forces of growth respectively, thus biasing the estimated completeness of reporting of deaths. Constants, $\delta_{t,i,s}$, are used in order to produce corrected forces (${}_{y-x}\tilde{r}_{i,s,x}$) and thus eliminate the bias as follows:

$$ {}_{y-x}\tilde{r}_{i,s,x} = {}_{y-x}r_{i,s,x} + \delta_{t,i,s} \qquad (5.21) $$

Then equation 5.15 can be rewritten as:

$$ \hat{N}_{t,i,s,x} = \hat{N}_{t,i,s,x+5} \exp(5.{}_5\tilde{r}_{i,s,x}) + {}_5D_{t,i,s,x}{}^R \exp(2{,}5.{}_5\tilde{r}_{i,s,x}) \qquad (5.22) $$

Note that the Preston and Coale method is a specific case of the more general Bennett and Horiuchi method, with ${}_5\tilde{r}_{i,s,x} = r$ for all values of s and x.

Bennett and Horiuchi suggested that the following equation be used to estimate the population aged X, where X is the lower bound of the open interval:

$$ \hat{N}_{t,i,s,X} = {}_{\omega-X}D_{t,i,s,X}{}^R \{\exp({}_{\omega-X}\tilde{r}_{i,s,X}.e_{i,s,X}) - ({}_{\omega-X}\tilde{r}_{i,s,X}.e_{i,s,X})^2 / 6\} \qquad (5.23) $$

where $e_{i,s,x}$ denotes the expectation of life for an individual in population group i with gender s and age x.

Equation 5.23 is used to calculate the starting value to be used in the iterative equation 5.22.

The derivation of equation 5.23 is covered an appendix of the Bennett and Horiuchi paper. Although the choice of suitable values of $e_{i,s,X}$ may be somewhat arbitrary, they show that the resulting estimates of completeness of reporting are not significantly biased.

Bennett and Horiuchi suggest that the following ratios be calculated:

$$c_{t,i,s,x}{}^{4} = {}_{10}\hat{N}_{t,i,s,x-5} / {}_{10}N_{t,i,s,x-5} \qquad (5.24)$$

As in the Preston and Coale method, cumulative ratios may be calculated as follows:

$$c_{t,i,s,x}{}^{5} = \sum_{a=x}^{(Y-5+4x)/5} {}_{5}\hat{N}_{t,i,s,5(a-x)+x} \Big/ \sum_{a=x}^{(Y-5+4x)/5} N_{t,i,s,5(a-x)+x} \text{ for some } Y \qquad (5.25)$$

A plot of the ratios shown in equations 5.24 and 5.25 against x should yield horizontal lines. $\delta_{t,i,s}$ can be adjusted in order to produce the most horizontal lines.

The Bennett and Horiuchi estimate of the completeness of reporting of deaths occurring during year t amongst individuals in population group i and with gender s ($C_{t,i,s}$) is calculated by taking the mean or median of the ratios $c_{t,i,s,x}{}^{4}$ or $c_{t,i,s,x}{}^{5}$ over the adult ages (e.g. ages 20 to 60).

Possible Distortions

The results of the Bennett and Horiuchi method may be distorted for the reasons set out in section 5.3.2.

5.3.4 Application of the Preston and Coale and the Bennett and Horiuchi Methods

The following data were used:

- ${}_{5}D_{t,4,s,x}{}^{R}$: Estimates of the number of reported black deaths in 1991, 1992, 1993 and 1994, assuming 100% (Scenario 1), 90% (Scenario 2) and 80% (Scenario 3) of the asian, coloured and white deaths were reported, as derived in section 5.3.1.
- ${}_{5}N_{t,4,s,x}$: Estimates of the black population at the mid-point of years 1991 to 1994. These population estimates are discussed in detail in section 5.4.

The lower bound of the open interval was chosen to be 65 (i.e. $X = 65$). A wider interval would be preferable to incorporate all age misreporting, although the estimate of the weight of the deaths in the interval would become less reliable. A narrower open interval is suggested in the literature. The chosen X was believed to be a satisfactory compromise. For the purposes of comparison, the Preston and Coale and the Bennett and Horiuchi methods were applied using an X of 75. The resulting curves were more erratic than those obtained using an X of 65, providing further justification for the chosen X.

The Preston and Coale method was applied as follows:

$\hat{N}_{t,4,s,65}$ was calculated as shown in equations 5.16 and 5.17 using estimates from the western pattern of Coale-Demeny life tables in equation 5.17. As a comparison, estimates of $z(65)$ were derived using equation 5.18, with values of $d_{i,s,x}{}^{S}$ taken from Dorrington's mortality table (Dorrington, Martens and Slawski, 1991) with the adjustments outlined in section 5.6.1. The estimates were similar to those derived using equation 5.17, thus justifying the use of the latter. The annual force of growth (r) was adjusted separately for each scenario in order to produce relatively horizontal curves of the ratios $c_{t,4,s,x}{}^{3}$ against x, for each t and s. A force of 2,9%, 2,8% and 3,0% was assumed for Scenarios 1, 2 and 3 respectively. These are fairly consistent with those used in previous applications of the Preston and Coale method (Dorrington: Personal communication).

The Bennett and Horiuchi method was applied as follows:

Annual age-specific forces of growth ($_{y\text{-}x}r_{4,s,x}$) for the black population were required for the period 1991 to 1994. Sadie's estimates of the black population in 1991 and 1996 (Sadie, 1993) were used as shown in equation 5.20, as the methodology and assumptions implicit in these estimates are likely to be consistent. Inconsistencies in methodology exists between the 1991 and (preliminary) 1996 population censuses conducted by the CSS. Although Sadie's population estimates may be incorrect (discussed in section 5.4.2), the age-specific forces of growth have not been seriously disputed.

Equation 5.23 was applied to the data, using:

$e_{4,0,65} = 11{,}21$ and
$e_{4,1,65} = 14{,}04$.

These values were calculated using the adjusted black mortality table derived by Dorrington, Martens and Slawski (1991), described in section 5.6.1.

The constants, $\delta_{t,4,s}$, were initially set equal to zero, and the population estimates calculated as shown in equation 5.22. The ratios $c_{t,4,S,x}{}^{4}$ and $c_{t,4,S,x}{}^{5}$ were calculated as shown in equations 5.24 and 5.25.

Values for $\delta_{t,4,s}$ were chosen so as to minimise the sum of the absolute deviations of the ratios $c_{t,4,s,x}{}^4$ from their mean, over the age range 20 to 60. The curves of the ratios $c_{t,4,s,x}{}^4$ and $c_{t,4,s,x}{}^5$ were thus made as horizontal as possible over these age ranges.

Results

Plots of the ratios $c_{t,4,s,x}{}^3$ (for the Preston and Coale method) and $c_{t,4,s,x}{}^5$ (for the Bennett and Horiuchi method) against x under the three scenarios, are shown in Figures 5.5 to 5.10.

For the Preston and Coale method, the completeness of reporting was defined as the mean of the ratios $c_{t,4,s,x}{}^3$ over ages 20 to 60, as shown in equation 5.26:

$$C_{t,4,s} = \frac{1}{8}\sum_{x=20}^{27} c_{t,4,s,5(x-20)+20}{}^3 \qquad (5.26)$$

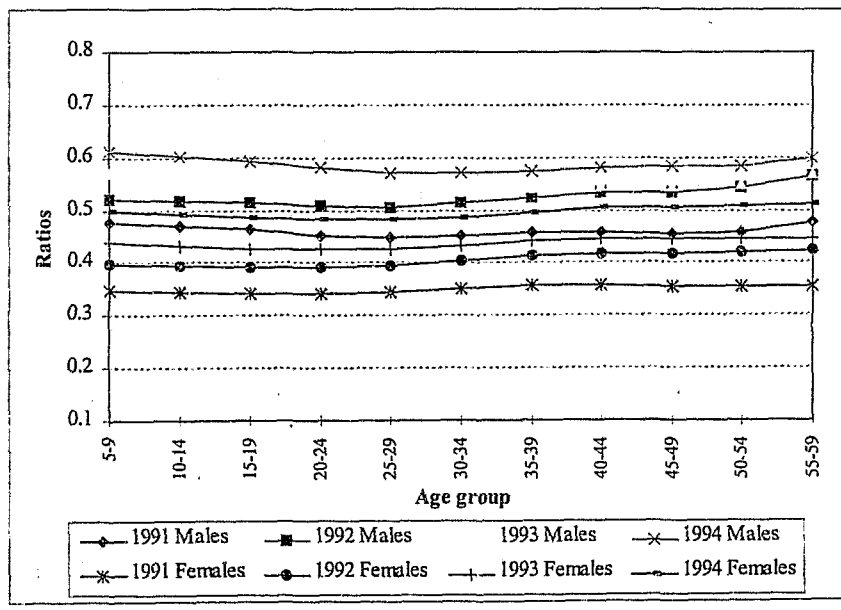


Figure 5.5 Scenario 1: Plot of $c_{t,4,s,x}{}^3$ against x for the black population using the Preston and Coale method

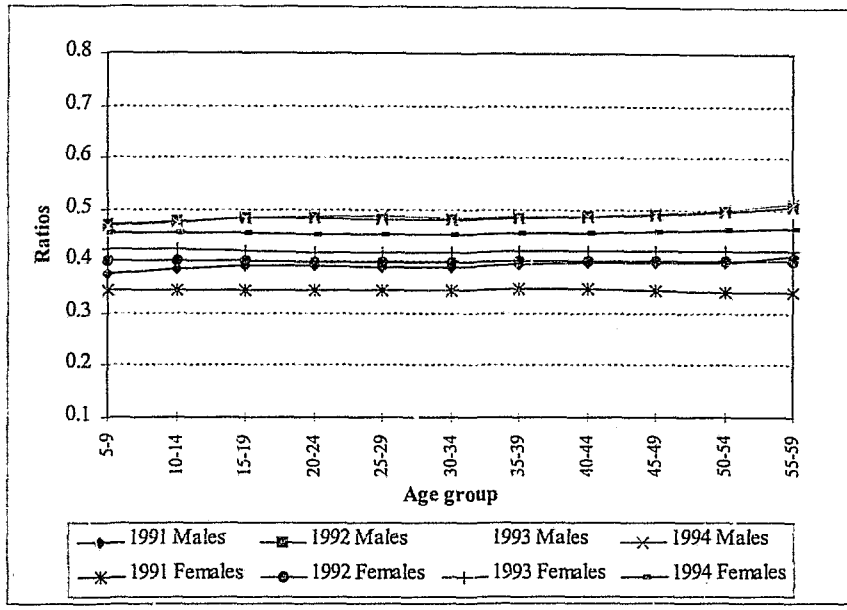


Figure 5.6 Scenario 1: Plot of $c_{t,4,s,x}{}^{5}$ against x for the black population using the Bennett and Horiuchi method

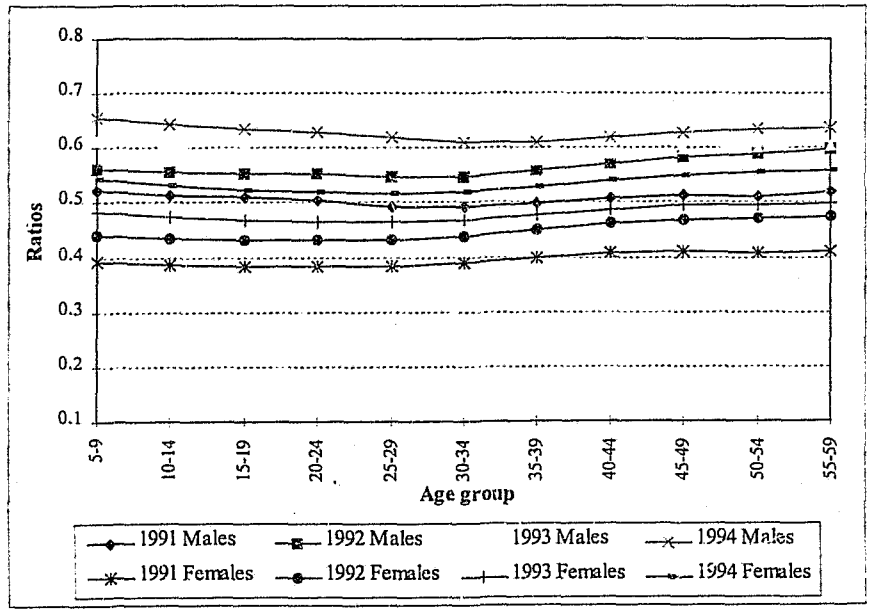


Figure 5.7 Scenario 2: Plot of $c_{t,4,s,x}{}^{3}$ against x for the black population using the Preston and Coale method

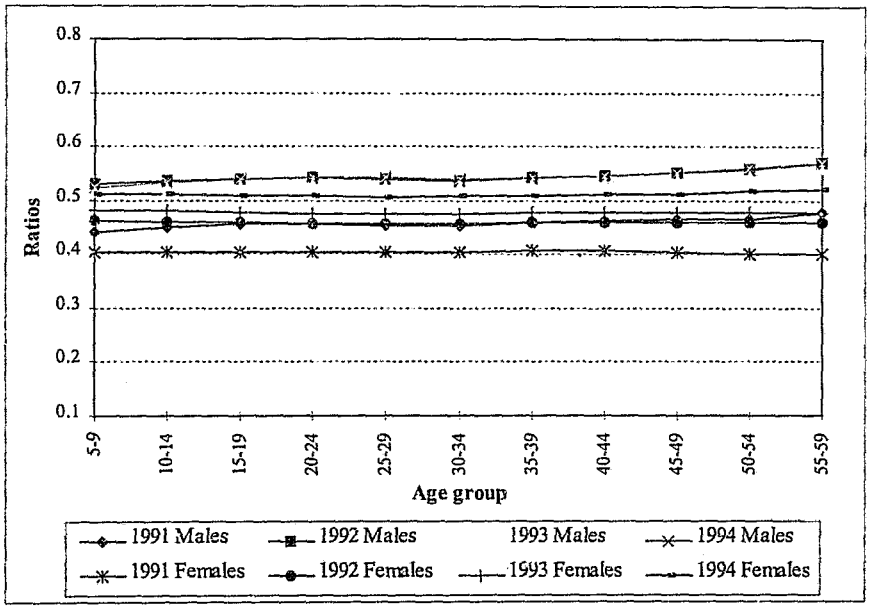


Figure 5.8 Scenario 2: Plot of $c_{t,4,s,x}{}^{5}$ against x for the black population using the Bennett and Horiuchi method

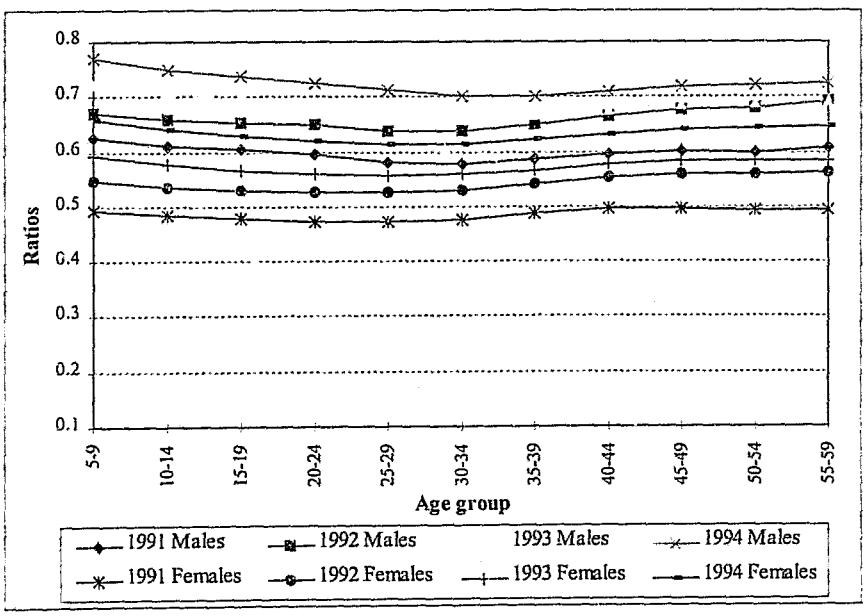


Figure 5.9 Scenario 3: Plot of $c_{t,4,s,x}{}^{3}$ against x for the black population using the Preston and Coale method

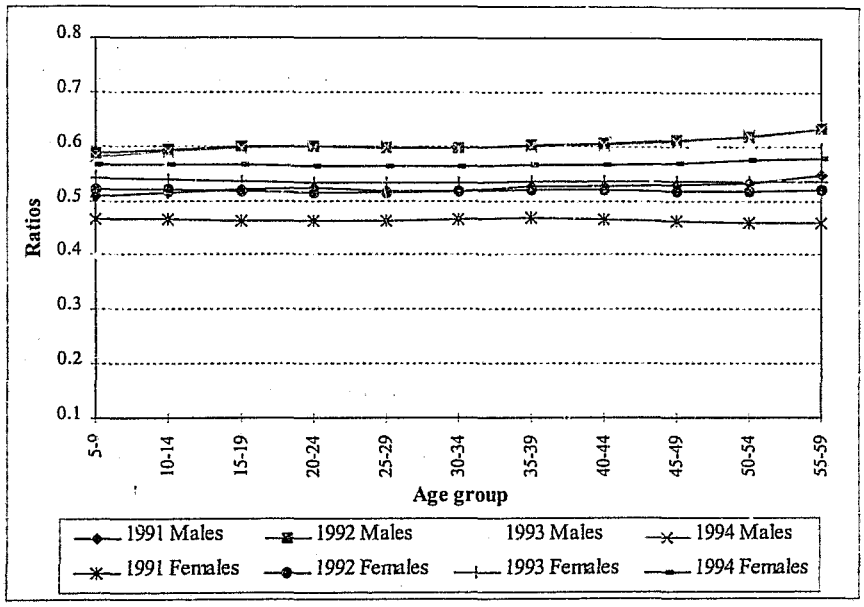


Figure 5.10 Scenario 3: Plot of $c_{l,A,s,x}{}^{5}$ against x for the black population using the Bennett and Horiuchi method

Table 5.3 Estimated completeness of reporting of black deaths using the Preston and Coale (P & C) and the Bennett and Horiuchi (B & H) methods

Year	Scenario 1		Scenario 2		Scenario 3	
	P & C	B & H	P & C	B & H	P & C	B & H
Males						
1991	45,6%	39,6%	50,4%	46,2%	59,3%	52,7%
1992	52,1%	48,8%	56,8%	54,8%	66,1%	60,8%
1993	54,3%	47,9%	58,9%	53,5%	67,9%	59,2%
1994	58,0%	49,2%	62,3%	54,9%	71,4%	60,7%
Females						
1991	35,1%	34,4%	39,9%	40,3%	48,6%	46,3%
1992	40,7%	40,0%	45,3%	45,8%	54,4%	51,7%
1993	43,6%	41,9%	48,1%	47,7%	57,2%	53,4%
1994	49,4%	45,6%	53,6%	51,2%	62,9%	56,7%

For the Bennett and Horiuchi method, the completeness of reporting was defined as the mean of the ratios ${}_5c_{t,4,s,x}$ over ages 20 to 60, as shown in equation 5.27:

$$C_{t,4,s} = \frac{1}{8}\sum_{x=20}^{27} {}_5c_{t,4,s,5(x-20)+20} \qquad (5.27)$$

The estimates of $C_{t,4,s}$ for the three scenarios and using both methods are shown in Table 5.3.

5.3.5 Comparison and Discussion of Results from the Preston and Coale and the Bennett and Horiuchi Methods

The following points are noted on examination of the results:

- A higher proportion of male deaths are reported than female deaths.
- The completeness of reporting generally increases from 1991 to 1994, although reduces slightly for males in 1993 when using the Bennett and Horiuchi method.
- The Preston and Coale method generally results in higher estimates of the completeness of reporting than the Bennett and Horiuchi method.
- As expected, Scenario 1 results in lower estimates of the completeness of death registration than Scenario 2, which in turn results in lower estimates than Scenario 3.

It has been noted (CSS, 1996 a) that the former KaNgwane and KwaNdebele experienced problems relating to the recording of deaths during 1991. Similar problems were experienced by the former KwaZulu, Gazankulu, QwaQwa and KwaNdebele during 1992. The recorded deaths for these years may thus be unusually low.

Deaths arising in the former TBVC states were excluded in the death reports until 1994. However, the population data used in the analysis include estimates of the population living in these areas. The relatively large increase in the proportion of reported black female deaths between 1993 and 1994 may thus reflect the inclusion of deaths arising in the former TBVC states during 1994.

On inspection of the curves in Figures 5.5 to 5.10, those derived using the Bennett and Horiuchi method appear flatter than those derived using the Preston and Coale method. This suggests that the assumption of a constant force of growth may be inaccurate for the black population, and that the age-specific forces used in the Bennett and Horiuchi method are more realistic. Empirical evidence appears to justify this. Black fertility rates have gradually declined over the past few decades (discussed in section 5.4.2), and black mortality is likely to have improved. The assumption of stability is thus not realistic for the black population, implying that

a constant force of growth in the Preston and Coale method may introduce distortion. This may overstate the estimates of the proportion of deaths reported, and would explain the discrepancy between the estimates from the two methods.

It is unlikely that 100% of asian, coloured and white deaths were reported in the years 1991 to 1994. The estimates under Scenario 1 are thus likely to be on the low side.

In order to comment on the reasonability of Scenarios 2 and 3, the results have been compared to estimates from independent sources.

The Preston and Coale method was applied by Dorrington (1989) and Dorrington, Martens and Slawski (1991) in order to estimate mortality rates for the black population. In his 1989 paper, Dorrington examined deaths reported during 1979, 1980 and 1981. It was estimated that 53,4% of male deaths and 43% of female deaths were reported during this period. In the 1991 paper, deaths reported around 1985 were examined. It was estimated that 54% of male deaths and 41% of female deaths were reported. These estimates would be distorted if the black population in 1980 and 1985 was not stable.

The Department of Health produced a fairly crude estimate for 1994 (personal communication) by dividing reported deaths by expected deaths derived using life tables produced by the US Census Office. 50,34% of deaths from all population groups were estimated to have been reported. Since the completeness of death reporting is lower for blacks than for asians, coloureds and whites, it is inferred that less than 50,34% of black deaths were reported in 1994.

Scenario 2 appears fairly consistent with the Department of Health's estimate. If Dorrington's estimates are accurate, the death registration for the black population appears to have deteriorated between 1985 and 1991. One may reason that such deterioration is feasible due to the political upheaval of the late 1980's. However, such reasoning remains speculative without further corroboration.

The 1991 estimates under Scenario 3 are more consistent with Dorrington's estimates. This scenario would be realistic if the completeness of death registration had increased slightly for females and reduced slightly for males between 1985 and 1991.

5.3.6 The Estimated Number of Actual Black Deaths

The estimates of the actual number of black deaths were calculated for Scenarios 2 and 3 only, for reasons discussed in section 5.3.5. The Bennett and Horiuchi method was used to obtain estimates of the completeness of deaths registration.

The actual number of deaths were calculated using equation 5.28:

$$ {}_5D_{t,4,s,x} = {}_5D_{t,4,s,x}{}^{R} / C_{t,4,s} \qquad (5.28) $$

Summarised results are shown in Table 5.4. The detailed age and sex structure of deaths under Scenario 2 is shown in Appendix B.

The following observations are made:

- Male deaths constituted between 55% and 59% of the total deaths in each year.
- Fewest deaths occurred during 1992. This is probably an artificial phenomenon, as deaths from 1991, 1993 and 1994 suggest an increasing trend.
- A greater number of deaths are estimated to have occurred under Scenario 2 than under Scenario 3.

Table 5.4 Estimated number of actual black deaths

Year	Total	Males	Females
Scenario 2			
1991	242 090	141 664	100 425
1992	207 481	116 025	91 456
1993	248 978	144 605	104 373
1994	257 664	150 065	107 599
Scenario 3			
1991	227 389	132 126	95 263
1992	199 984	111 783	88 201
1993	238 401	138 235	100 166
1994	247 143	143 295	103 847

5.4 Estimation of the Size and Demographic Structure of the Black Population

5.4.1 Available Data

South African population data are subject to a great deal of uncertainty. Black population statistics are especially unreliable.

The CSS conducted a population census in 1991, excluding the population living in the former TBVC states. Aerial photographs were used to estimate the size of populations living in informal settlements. Certain parts of the country (with the majority of residents being black) were not clearly demarcated into enumerator areas, and enumerators were left to establish a population size as best they could (CSS, 1997).

The CSS adjusted the census count using Sadie's population projection model (CSS, 1991 b). The black population was projected from the 1970 census data which were believed to be fairly accurate. Fertility and mortality rates were estimated, and net migration was assumed to be negligible. The under-count for blacks was found to be 16,8%.

The total South African population was put at 38,0 million in 1991. Projections based on the 1991 census results suggest a population size of 42,1 million in 1996 (CSS, 1996 a and CSS, 1997).

The CSS has published preliminary estimates of the 1996 census (CSS, 1997). Information on the distribution of the population by population group, gender and province is not available as yet. The preliminary results indicate a population size of 37,9 million, which is substantially lower than previously estimated.

5.4.2 Comparison of the 1991 and Preliminary 1996 Census Results

Arguments in Support of the Preliminary 1996 Census Results

Fertility Rates. It has been suggested that the fertility rates used to adjust the 1991 census results were too high, particularly for the black population (CSS, 1997).

Total fertility rates for black women assumed in Sadie's model (CSS, 1991 b) are shown in Figure 5.11. Sadie has since revised his fertility estimates downwards (Sadie, 1993).

In a recent paper by Udjo (1997), the birth histories from the 1995 October Household Survey (CSS, 1996 b) were examined. A model proposed by Brass (1974, 1981) and developed by Booth (1979) and Zaba (1981) was fitted to the data. This model was designed to estimate the distribution of births from reports on births during the previous twelve months. Total fertility rates for the whole population were estimated, and are illustrated in Figure 5.12.

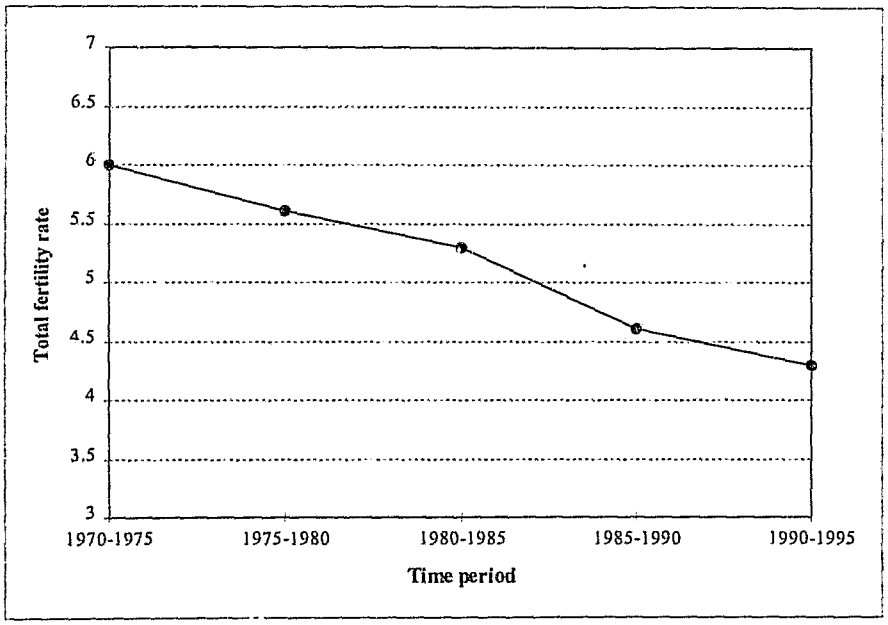


Figure 5.11 Total fertility rates for black women used in Sadie's model

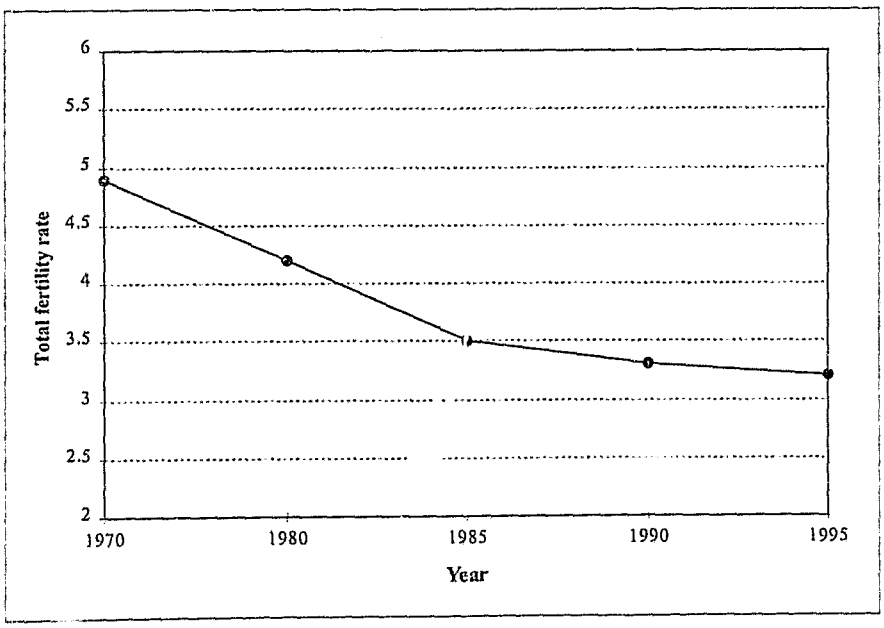


Figure 5.12 Udjo's estimated total fertility rates

Udjo applied these fertility rates to the 1970 census data (together with mortality rates estimated from the 1995 October Household Survey). Three estimates of the 1996 population were obtained: 37 494 000 (low variant), 37 990 000 (medium variant) and 39 165 000 (high variant). The low variant is based on the unadjusted 1970 census results, the medium variant is based on the 1970 census results with adjustments made by Sadie (1973), and the high variant is based on the 1970 census results with adjustments made by the CSS. The low and medium variants fall within the 95% confidence interval for the preliminary 1996 census results.

Mortality rates, relative to fertility rates, do not significantly affect population projections. This implies that the 1996 population as estimated by Udjo should not differ significantly from that estimated by Sadie as a result of differing mortality assumptions, but rather due to differing fertility assumptions.

An estimate of the total fertility rate for the black population for the period 1990 to 1995 can be calculated from Udjo's corresponding rate for the whole population, as follows:

- Suppose that the total fertility rate for the whole population for the period 1990 to 1995 was equal to Udjo's estimate for 1990 (i.e. 3,3).
- Let TFR_i denote the total fertility rate for population group i.
- Let p_i denote the average proportion of all women aged between 15 and 50 that are in population group i for the period 1990 to 1995.

Then $\sum_{i=1}^{4} TFR_i . p_i = 3{,}3$

TFR_4 can be calculated assuming values for p_i, TFR_1, TFR_2 and TFR_3.

Suppose that the fertility assumptions used to adjust the asian, coloured and white population groups in the 1991 census are valid. Then, $TFR_1 = 2{,}45$; TFR_2 [illegible] and $TFR_3 = 1{,}81$ (CSS, 1991 b).

Suppose that $p_1 = 3\%$, $p_2 = 10\%$, $p_3 = 14\%$ and $p_4 = 73\%$. These proportions are fairly consistent with the black population estimates derived at the end of this section, and the asian, coloured and white population estimates derived in section 5.3.1.

$TFR_4 = 3{,}7$.

This is lower than the black fertility rate used in Sadie's model.

Urbanisation and its Effect on Fertility Rates. The preliminary 1996 census results suggest that the South African population is more urbanised than expected.

The preliminary census results estimate that 55,4% of the population was urbanised in 1996. This is a notably larger percentage than previous estimates of 49,7% (CSS, 1996 b) and 52% (Calitz, 1996).

The findings of a recent study conducted by the Human Sciences Research Council (Mostert and Hofmeyer, in press) suggest that fertility rates amongst black women are lower in the "urbanised provinces" (e.g. Gauteng, Kwazulu-Natal, Western Cape) than in the "rural provinces". Fertility rates were found to be particularly high in provinces incorporating the former homelands and TBVC states (e.g. Eastern Cape, Northern Province, Free State).

This theory is corroborated by a study conducted in the former Bophuthatswana (Directorate of Statistics, undated) where the total fertility rate in urban areas was found to be 3,08 compared to 4,16 in rural areas.

A higher percentage of the population living in urban areas would thus result in lower fertility rates. If the urbanisation statistics released by the CSS are accurate, it is likely that previous fertility estimates, including those used to adjust the 1991 census, were too high.

Arguments Against the Preliminary 1996 Census Estimates

A number of criticisms of the preliminary 1996 census estimates have emerged. It has been suggested that the preliminary estimates understate the true population size.

Comparison with Other Population Projections. Critics feel it is very unlikely that previous estimates could have overstated the size of the population to such a large extent (10%). The size of the 1996 population was estimated by the Development Bank of Southern Africa to be 44,4m (Calitz: Personal communication). This is even greater than the projections made by the CSS based on the 1991 census results.

Male:Female Ratio. Young populations (i.e. populations with a larger than average proportion of young people) tend to consist of a similar number of males and females. It has been suggested that the South African population is too young to contain 1,5m more females than males, as is indicated by the preliminary census results.

Females are estimated to have constituted 47,2% of the population in 1970 (CSS, 1970). It is almost impossible, within the boundaries of reasonable assumptions, for this percentage to have increased by 4,8% to 52,0% in only 26 years (Dorrington: Personal communication).

The preliminary census results are thus disputed.

Comparison with the 1994 Election Statistics. Schlemmer and Levitz (1997) estimated the number of individuals aged 18 and older from the preliminary census results. This figure was compared to the number of voters in the 1994 election, based on the official figures of the Independent Electoral Commission. Since the preliminary estimates do not indicate the age distribution of the population, the age distribution was estimated using the 1995 October Household Survey and estimates from the Development Bank of Southern Africa.

The results of this comparison follow:

• Estimated population aged 18 and older in 1996:	21,7m
• Number of votes cast in 1994 election:	19,7m
• Percentage poll based on above:	91%
• Percentage poll if number of voters is increased by 1,9% per annum from 1994 to 1996:	94%

The authors suspect that a poll of more than 90% is highly unlikely in South Africa, particularly since such a high poll is unusual elsewhere in the world. The enthusiasm of South Africa's first democratic election may, however, have resulted in such a high poll.

The authors note that if the assumed age distribution is incorrect, the error would understate, rather than overstate, the population aged 18 and over, since fertility rates are believed to have fallen. The estimated poll is thus a lower bound for the actual poll.

Comparison with Statistics on Scholars. Schlemmer and Levitz (1997) estimated the number of children aged 6 to 14 from the preliminary 1996 census estimates. This figure was compared to the number of scholars aged 6 to 14 in 1995 as compiled by the Directorate of Information from the Department of Education. The age distribution of the population was estimated using the 1995 October Household Survey and estimates from the Development Bank of Southern Africa.

The results of this comparison follow:

• Estimated number of children aged 6 to 14 in 1996:	7,9m
• Number of scholars aged 6 to 14 in 1995:	8,37m
• Estimated percentage of children aged 6 to 14 attending school based on above:	106%
• Estimated percentage of children aged 6 to 14 attending school assuming school attendance increased by 1,5% between 1995 and 1996	108%

A school attendance figure of more than 100% appears to be nonsensical. However, a number of scholars are unusually old for their classes, while some younger children may have started school earlier. The figures produced by the Department of Education may thus be overstated if age was measured with reference to the class attended.

Nevertheless, the school attendance figure is likely to be a lower bound for the actual school attendance figure since the assumed age distribution may in fact overstate the number of children in the population. The authors are thus of the opinion that the preliminary results are a very serious underestimate of the true population size.

Resolution

It was decided to base estimates of the black population on the 1991 census results. Although the preliminary 1996 results are allegedly based on empirical evidence, the current lack of detail restricts closer examination. The rejection of the fully published 1991 results in favour of the preliminary 1996 results can therefore not be justified. Moreover, differences in the enumeration of individuals aged less than 20 are not important in the estimation of AIDS-related deaths.

5.4.3 Method Used to Estimate the Black Population

The census estimates are grouped into the quinquennial age bands (0, 5), (5, 10), (10, 15), etc.. Thus, in the notation defined in section 5.2, x is a multiple of 5, and y-x equals 5. The population will only be estimated up to and including the age group (70, 75).

The 1991 census estimates of the black population (adjusted for undercount) did not include the population living in the former TBVC states. The population in these TBVC states thus had to be estimated and added to the census figures in order to accurately estimate the total black population of South Africa.

The 1991 census estimates were assumed to represent the population at the mid-point of the year. Although the census was conducted on 7th March, the inaccuracy was assumed to be negligible in comparison to the uncertainty surrounding the figures themselves.

${}_{s}N_{1,4,s,x}^{PROV}$ was defined as the 1991 census estimates of the black population living in the former provinces (including the former self-governing territories) with gender s and aged between x and x+5.

The CSS estimated that the population of the TBVC states on 7th March 1991 was 6,751 million (CSS, 1993 a). However, no age or gender distribution was given. Calitz (1991), under the auspices of the Development Bank of Southern Africa

(DBSA), estimated the TBVC population in 1995 with a breakdown by gender and age. Broad age bands were used: (0, 15), (15, 20), (20, 65), (65, ω). The CSS estimate of 6,751 million was split into these broad bands, and then allocated further into the quinquennial age bands using the age-distribution of blacks living in the provinces or self-governing territories.

The number of black individuals living in the former TBVC states with gender s and aged between x and $x+5$ as at 7th March 1991 was calculated as follows:

$$ {}_5N_{1,4,s,x}^{TBVC} = 6751000\left(\frac{{}_{b-a}N_{5,4,s,a}^{DBSA}}{\sum_{s=0}^{1}{}_{\omega}N_{5,4,s,0}^{DBSA}}\right)\left(\frac{{}_5N_{1,4,s,x}^{PROV}}{{}_{h-a}N_{1,4,s,a}^{PROV}}\right) \quad \text{for } x \geq a \text{ and } y \leq b \tag{5.29} $$

where a and b are the lower and upper bounds of a particular age group used in the DBSA estimates, and ${}_{y-x}N_{t,4,s,x}^{DBSA}$ denotes the number of blacks living in the former TBVC states with gender s and aged between x and y on 30th June of year t, as estimated by the Development Bank of Southern Africa (Calitz, 1991).

The final estimate of the total South African black population on 30th June 1991 was calculated as:

$$ {}_5N_{1,4,s,x} = {}_5N_{1,4,s,x}^{PROV} + {}_5N_{1,4,s,x}^{TBVC} \tag{5.30} $$

The black population at the mid-point of year $t+1$ ($t = 1, 2, 3$) was projected from the 1991 estimates for quinquennial age groups up to age 75 as shown in equation 5.31:

$$ {}_5N_{t+1,4,s,x+1} = {}_5N_{t,4,s,x} - {}_5D_{t,4,s,x}^{R} / C_{t,4,s} \tag{5.31} $$

where $C_{t,4,s}$ is the Bennett and Horiuchi estimate of the completeness of reporting of black deaths with gender s in year t.

Different scenarios will result in different population estimates, since $C_{t,4,s}$ is dependent on the scenario chosen. Population estimates were calculated for Scenarios 2 and 3 in order to correspond with the estimated actual black deaths (shown in Table 5.4).

Equation 5.31 can only be applied for year $t+1$ once the completeness of reporting has been estimated for year t, which in turn relies upon the population estimate for year t.

Before projecting for the following year, the projected population was regrouped

into the original quinquennial age bands to correspond to the death data, as indicated in equation 5.32:

$$
\begin{aligned}
{}_5N_{t+1,4,s,x} &= {}_5N_{t+1,4,s,x+1}\left(\frac{{}_4l^{s}_{4,s,x+1}}{{}_5l^{s}_{4,s,x+1}}\right) + {}_5N_{t+1,4,s,x-4}\left(\frac{{}_5l^{s}_{4,s,x}}{{}_5l^{s}_{4,s,x-4}}\right) \quad \text{for } x = 5, 10, \ldots, 70 \\
{}_5N_{t+1,4,s,x} &= {}_5N_{t+1,4,s,x+1}\left(\frac{{}_4l^{s}_{4,s,x+1}}{{}_5l^{s}_{4,s,x+1}}\right) + B_{t+1,4,s} \quad \text{for } x = 0 \qquad (5.32)
\end{aligned}
$$

where values of ${}_{y-x}l^{s}_{4,s,x}$ are taken from Dorrington, Martens and Slawski's mortality table (1991) with the adjustments outlined in section 5.6.1.

As discussed in section 5.3.1, the proportion of births which are male is estimated to be less than 52% for the black population. Udjo (1997) suggests that 51,5% of black births are male. This is consistent with the ratio assumed for other countries south of the Sahara. $B_{t,4,s}$ was thus calculated as shown in equation 5.33:

$$
\begin{aligned}
B_{t,4,s} &= 0{,}515\sum_{x=15}^{21} {}_5N_{t,4,1,5(x-15)+15} \cdot {}_5f_{t,4,5(x-15)+15} \quad \text{for males} \\
B_{t,4,s} &= 0{,}485\sum_{x=15}^{21} {}_5N_{t,4,1,5(x-15)+15} \cdot {}_5f_{t,4,5(x-15)+15} \quad \text{for females} \qquad (5.33)
\end{aligned}
$$

${}_5f_{t,4,x}$ was calculated by spreading the estimated total fertility rate for blacks (section 5.4.2) across the relevant quinquennial age bands according to the fertility rates used to adjust the 1991 census as shown in equation 5.34:

$$
{}_5f_{t,4,x} = 3{,}7\frac{{}_5f_{4,x}^{CSS}}{\sum_{x=15}^{21} {}_5f_{4,5(x-15)+15}^{CSS}} \qquad \text{for all } t \qquad (5.34)
$$

where ${}_5f_{4,x}^{CSS}$ denotes the fertility rate for black women aged between x and x+5 which was used to adjust the 1991 census results (CSS, 1991 b).

5.4.4 Results

Summarised population estimates appear in Table 5.5, and detailed estimates under Scenario 2 in Appendix C.

The following observations are made:

- Females constitute approximately 51% of the population in each year. This is slightly lower than the 52% estimated in the Preliminary 1996 Census results.

The population estimates thus appear to be more reasonable than the Preliminary 1996 Census results, for reasons discussed in section 5.4.2.

- The population increased at an approximate rate of 2,3% per annum.
- The population estimates under Scenario 3 are larger than those under Scenario 2. This is expected, as fewer deaths are estimated to have occurred under Scenario 3 (Table 5.4).

Table 5.5 Estimated size of the black population aged 0 to 75

Year	Total	Males	Females
Scenario 2			
1991	28 088 782	13 690 266	14 398 516
1992	28 702 083	13 993 795	14 708 288
1993	29 376 056	14 335 818	15 040 238
1994	30 034 471	14 662 519	15 371 952
Scenario 3			
1991	28 088 782	13 690 266	14 398 516
1992	28 717 482	14 003 517	14 713 964
1993	29 399 760	14 350 136	15 049 624
1994	30 069 674	14 683 654	15 386 020

5.4.5 Possible Distortions

The population estimates are subject to uncertainty for the following reasons:

- The 1991 census results may not be an accurate representation of the black population, as suggested by the preliminary 1996 census results (CSS, 1997).
- The estimate of the size of the population living in the former TBVC states may be inaccurate.
- The assumed age distribution for the population living in the former TBVC states may be incorrect.
- Estimates of the number of babies born may be incorrect due to inaccurate fertility assumptions.
- Projections are dependent on the estimated actual black deaths, and will thus be affected by distortions in these estimates.
- Migration has been ignored.

5.5 Estimation of Actual Mortality Rates for the Black Population

The method used to calculate the actual average probabilities for year t that a black individual with gender s and aged between x and x+5 will die within one year (${}_5q_{t,4,s,x}{}^A$) is shown in equation 5.35:

$$ {}_5q_{t,4,s,x}{}^A = \frac{{}_5D_{t,4,s,x}}{{}_5N_{t,4,s,x} + 1/2 \cdot {}_5D_{t,4,s,x}} \qquad (5.35) $$

The derivation of equation 5.35 can be found in Neill (1977: 209-210).

It should be noted that the estimated actual mortality rate is subject to a great deal of uncertainty, as both the numerator and denominator of equation 5.35 may involve some distortion.

The actual mortality curves for 1991 to 1994 are indicated in Figures 5.13 to 5.16.

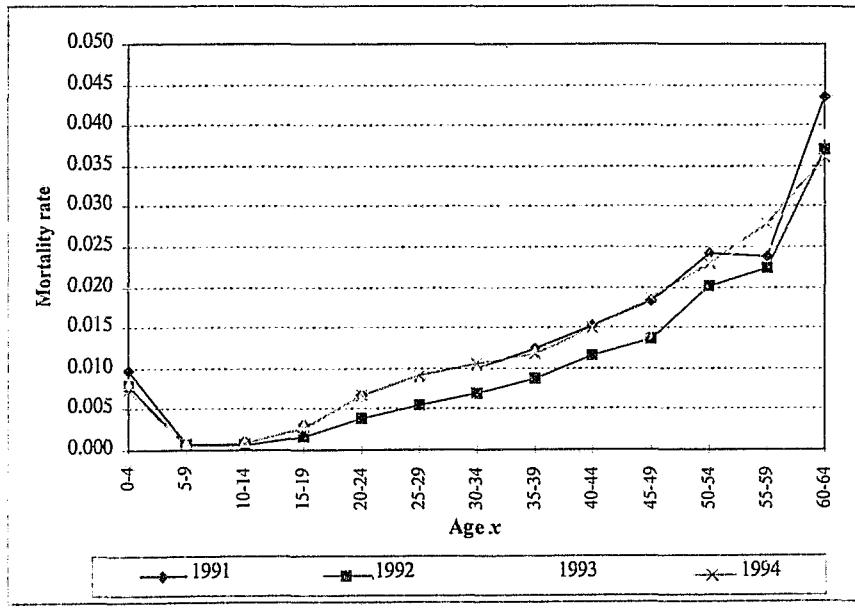


Figure 5.13 Estimated actual mortality rates for black males under Scenario 2

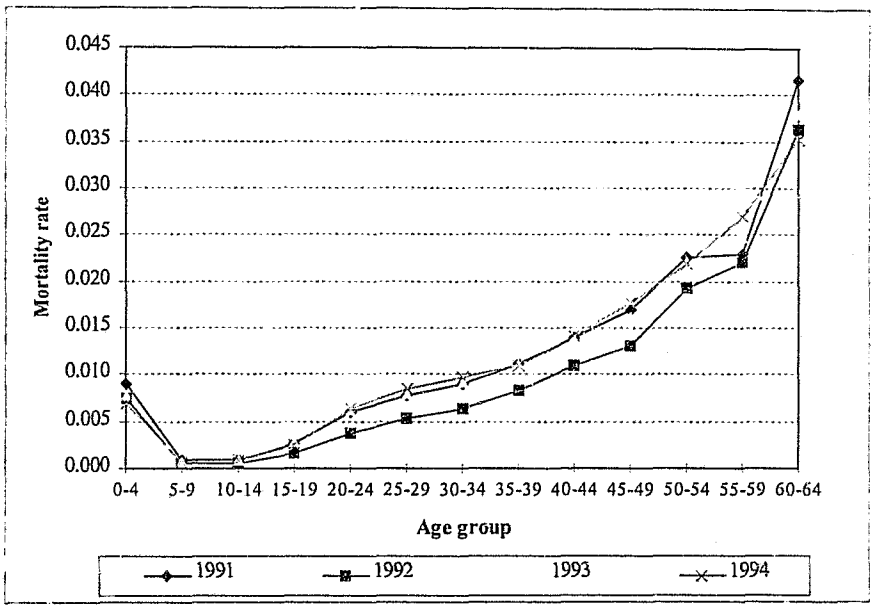


Figure 5.14 Estimated actual mortality rates for black females under Scenario 2

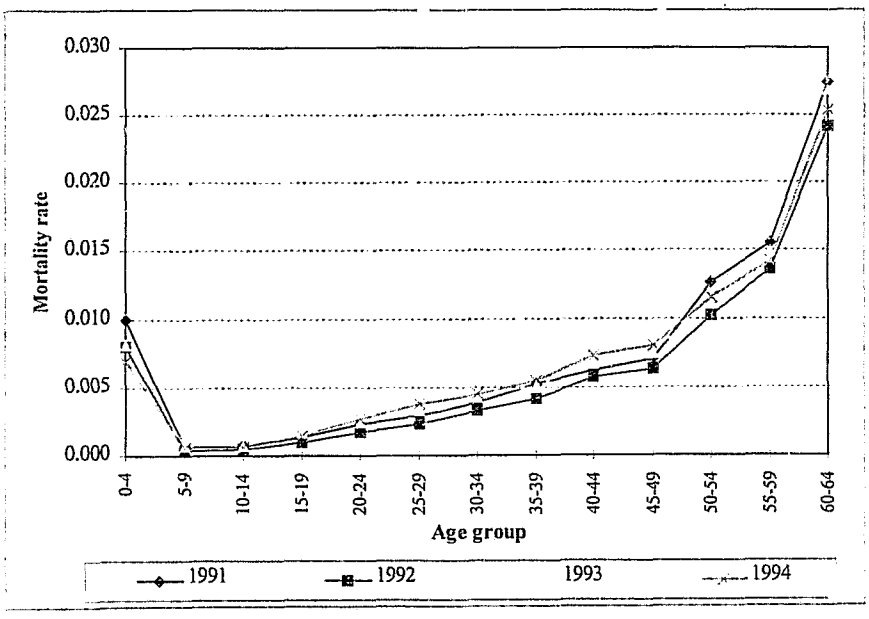


Figure 5.15 Estimated actual mortality rates for black males under Scenario 3

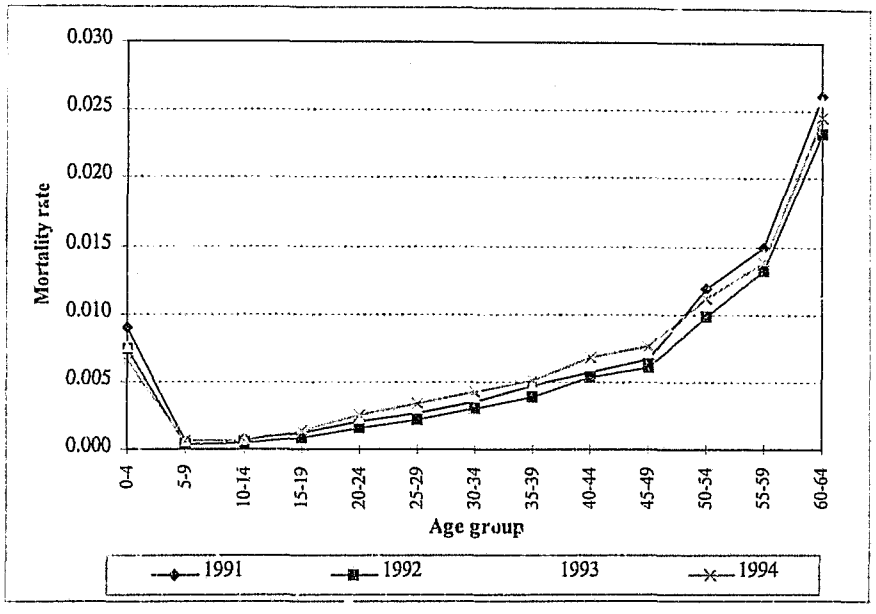


Figure 5.16 Estimated actual mortality rates for black females under Scenario 3

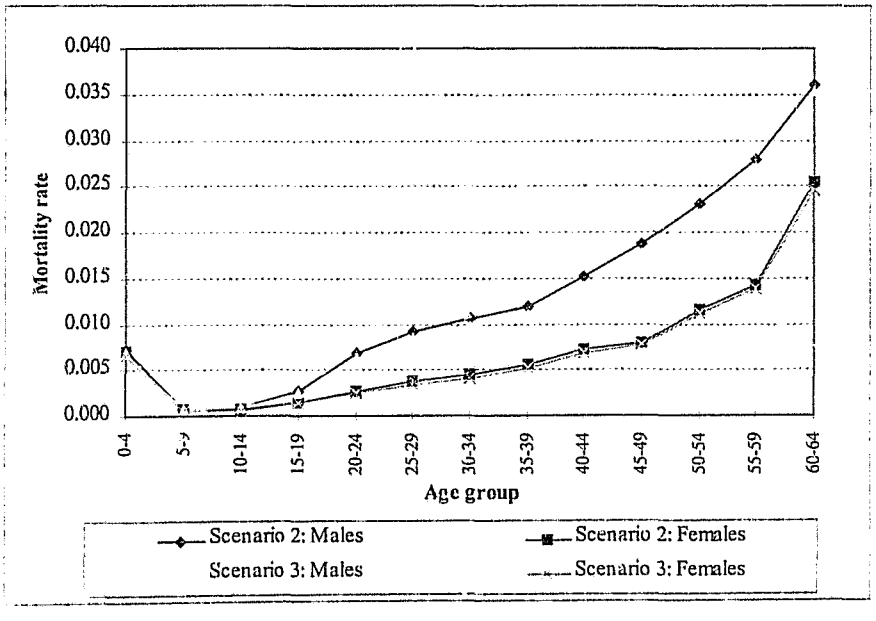


Figure 5.17 Estimated actual mortality curves for the black population in 1994

The following observations are made:

- Females experienced lower mortality than males at most ages, as expected.
- The gender-specific shapes of the actual mortality curves are similar in each year.
- The 1991 mortality rates are erratic at the older ages.
- Both genders experienced their lowest mortality during 1992. This is probably due to distortions in the underlying data apparent in Table 5.4. Female mortality appears to have worsened annually since 1992.

A comparison of the actual mortality curves for 1994 under Scenarios 2 and 3 is shown in Figure 5.17. From this figure, it is evident that the two scenarios result in mortality curves with similar gender-specific shapes, but with slightly different levels.

5.6 Selection and Adjustment of Expected Mortality Rates for the Black Population

5.6.1 Selection

AIDS-related deaths will be estimated by comparing the curve of actual mortality rates to a curve of expected mortality rates. The selection of expected mortality rates is thus an important process, as it directly affects the estimated number of such deaths. In order to produce accurate estimates, the expected mortality rate in year t applicable to black lives with gender s and aged between x and y (which will be denoted ${}_{y-x}q_{t,A,s,x}{}^{E}$) must equal the actual mortality rate from all causes except AIDS experienced by this same group of lives. However, these rates are not accurately known, and the expected mortality curve must be approximated.

Since substantial AIDS-related deaths have only recently emerged, mortality rates relating to a prior time should not be affected by AIDS. The disadvantage of using such rates is that they are likely to be out of date. It is possible to correct the level of the expected mortality curve, as discussed in section 5.6.2. However, if the shape of the mortality curve has changed over time, such historical rates would not be appropriate.

A further problem in the selection of suitable expected mortality rates is that very few life tables have been published for the black population in South Africa. The CSS justifies this by reasoning that the registration of vital events amongst the black population is incomplete and would thus lead to unreliable estimates (CSS, 1987), as has been shown in section 5.3.

The suitability of tables from populations with similar characteristics to the South

African black population has thus been assessed.

A number of mortality tables were considered for use as an expected mortality table:

Mortality Tables Produced by Dorrington

Dorrington produced two complete tables of estimated black mortality rates applicable to South Africa (Dorrington, 1989; Dorrington, Martens and Slawski, 1991). The Preston and Coale method (section 5.3.2) was used in the construction of both tables.

The earlier table relates to the period 1979-81, and relies upon 1980 census data (CSS, 1982) for the population estimates. The later table relates to a period centred around 1985, and relies on Sadie's population estimates (Sadie, 1988). The authors believed that Sadie's estimates produced the more convincing results, and the revised mortality rates were an improvement over Dorrington's previous estimates.

Both of the abovementioned tables needed to be extended to ages over 90. The rates are expected to be higher than those of the white, coloured and asian population groups, for reasons discussed in section 5.3.1. The rates were thus adjusted.

The mortality rates were adjusted as follows for males:

$$\begin{array}{ll} {}_{4,0,x}S \text{ unchanged} & \text{for } x < 76 \\ {}_{4,0,x}S = {}_{3,0,x}S + m_{4,0}(100\text{-}x) & \text{for } x = 76, 77, \ldots, 100 \\ {}_{4,0,x}S = {}_{3,0,x}S & \text{for } x > 100 \end{array} \qquad (5.36)$$

where $m_{4,0} = ({}_{4,0,75}S - {}_{3,0,75}S)/25$

The mortality rates were adjusted as follows for females:

$$\begin{array}{ll} {}_{4,1,x}S \text{ unchanged} & \text{for } x < 76 \\ {}_{4,1,x}S = {}_{3,1,x}S + m_{4,1}(100\text{-}x) & \text{for } x = 76, 77, \ldots, 100 \\ {}_{4,1,x}S = {}_{3,1,x}S & \text{for } x > 100 \end{array} \qquad (5.37)$$

where $m_{4,1} = ({}_{4,1,75}S - {}_{3,1,75}S)/25$

UN Life Tables for Developing Countries

The United Nations has constructed life tables for developing countries based on a number of patterns (UN, 1982). The general pattern is very similar to the Coale-Demeny west region (Coale and Demeny, 1966). The table was constructed using the eight-parameter formula developed by Heligman and Pollard (1980).

Zimbabwean Life Tables

The Central Statistical Office of Zimbabwe (1985) produced a life table for 1982.

Life Tables for Botswana

Udjo (1995) constructed various mortality tables for Botswana.

The life tables for Zimbabwe and Botswana are incomplete and thus not suitable for use. The UN mortality rates and those derived by Dorrington are plotted with 1991 estimated actual mortality rates under Scenario 2 in Figure 5.18.

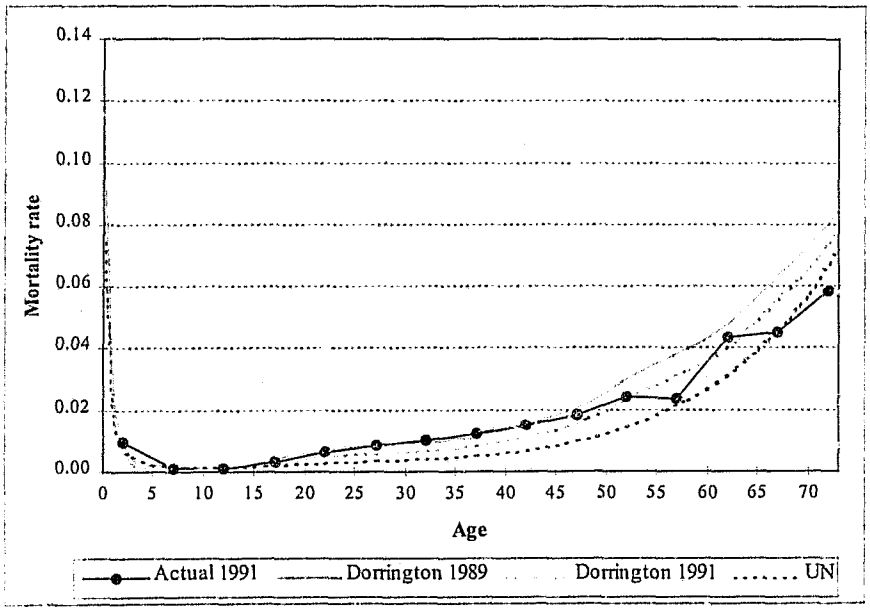


Figure 5.18 Comparison of various male African mortality rates and the 1991 actual rates under Scenario 2

The 1991 actual mortality rates become very erratic from age 55 onwards. This is unlikely to be a real feature of the black mortality schedule, but rather a result of distortions in the death and population statistics at these ages. The 1991 actual rates appear to converge towards Dorrington's 1991 rates up until age 50. The UN rates are generally lower than the 1991 actual rates.

It was thus decided to set the expected rates equal to the 1991 actual black mortality rates up until age 55, and Dorrington's 1991 rates thereafter. The curves of these expected rates under Scenarios 2 and 3 are shown in Figure 5.19. The distortions in the 1991 actual rates have been eliminated, and the rates increase relatively smoothly with age.

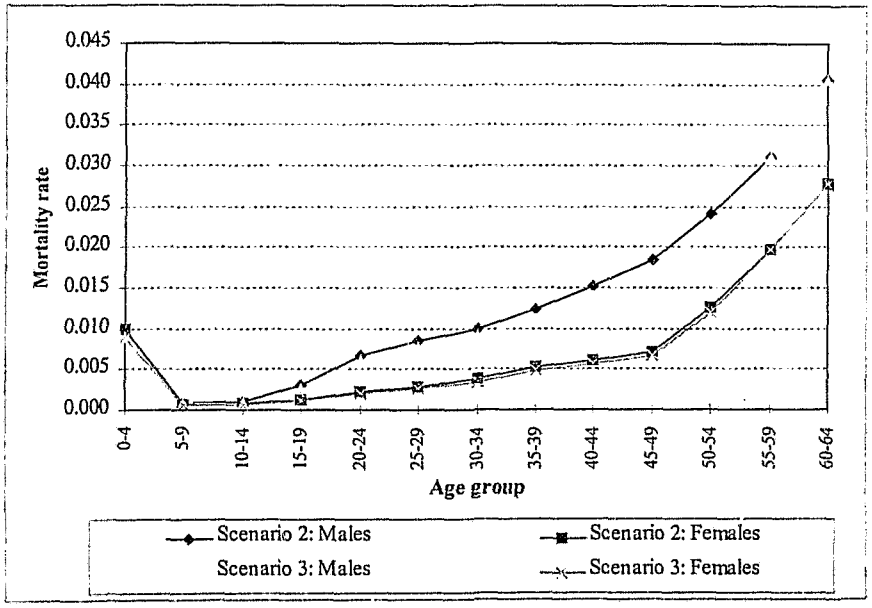


Figure 5.19 Expected mortality rates for the black population

It was assumed that no AIDS-related deaths occurred before 1992, and that a worsening of mortality since 1991 was the results of AIDS. If these 1991 rates do in fact incorporate AIDS mortality, the calculation of AIDS-related deaths in subsequent years would underestimate the true number of such deaths.

5.6.2 Adjustment

As mentioned briefly in section 5.6.1, the chosen expected mortality curve requires adjustment before use. AIDS is expected to alter the shape of the

mortality curve at the relevant ages (roughly from 20 to 60), but in the years being measured, it ought not to have a significant effect on the overall level of mortality. Changes in the level of mortality must thus be isolated from changes in the shape of the mortality curve which could arise due to AIDS. Moreover, reporting errors may create artificial differences in the level of the mortality curve. This section attempts to remove any changes in the overall level of the mortality curve in order to estimate AIDS-related deaths.

The proportion of expected deaths in year t amongst black individuals with gender s which occurred in the age group x to $x+5$ (denoted ${}_5w_{t,4,s,x}$) was calculated as follows:

$$ {}_5w_{t,4,s,x} = \frac{{}_5q_{t,4,s,x}^{E} \cdot {}_5N_{t,4,s,x}}{\sum_{a=0}^{14} {}_5q_{t,4,s,5a}^{E} \cdot {}_5N_{t,4,s,5a}} \qquad (5.38) $$

Note that $\sum_{x=0}^{14} {}_5w_{t,4,s,5x} = 1$.

The ratio of the actual to expected average mortality rates in year t for black individuals with gender s and aged between x and $x+5$ (denoted ${}_5R_{t,4,s,x}$) was calculated as follows:

$$ {}_5R_{t,4,s,x} = \frac{{}_5q_{t,4,s,x}^{A}}{{}_5q_{t,4,s,x}^{E}} \qquad (5.39) $$

The mean ratio of the actual to expected mortality rates in year t for black individuals with gender s was defined as:

$$ R_{t,4,s} = \sum_{x=0}^{14} {}_5R_{t,4,s,5x} \cdot {}_5w_{t,4,s,5x} \qquad (5.40) $$

If the actual and expected overall levels of mortality are the same, then $R_{t,4,S}$ will equal 1. However, if the actual level of mortality has improved at all ages, then $R_{t,4,S}$ will be less than 1. Conversely, $R_{t,4,S}$ will be greater than 1 if the actual level of mortality has worsened at all ages.

Adjusted expected mortality rates are defined as follows:

$$ {}_5\tilde{q}_{t,4,s,x}^{E} = {}_5q_{t,4,s,x}^{E} \cdot R_{t,4,s} \qquad (5.41) $$

The adjusted ratios of the actual to expected average mortality rates in year t for

black individuals with gender s and aged between x and x+5 (denoted by ${}_5\widetilde{R}_{t,4,s,x}$) are then defined as:

$$ {}_5\widetilde{R}_{t,4,s,x} = \frac{{}_5q_{t,4,s,x}^{A}}{{}_5\widetilde{q}_{t,4,s,x}^{E}} \qquad (5.42) $$

and the adjusted mean ratio of the actual to expected mortality rates in year t for individuals with gender s is defined as:

$$ \widetilde{R}_{t,4,s} = \sum_{a=0}^{14} {}_5\widetilde{R}_{t,4,s,5a} \cdot {}_5w_{t,4,s,5a} = 1 \qquad (5.43) $$

The mean ratio of the actual to expected mortality rates is thus forced to equal 1 by adjusting the expected mortality rates. Note that if actual mortality has improved (i.e. $R_{t,4,s}$ is less than 1), the expected mortality rates are reduced, thus lowering the overall level of the expected mortality curve. Conversely, if actual mortality has worsened (i.e. $R_{t,4,s}$ is greater than 1), the level of the expected mortality curve will be raised. In this way, the effect of changes in the level of the mortality curve is eliminated.

Implementation

The estimated values of $R_{t,4,s}$ are shown in Table 5.6. The values can be justified with reference to Figures 5.13 to 5.16. For example, it is clear that the expected mortality curve needs to be adjusted downwards for 1992, as the curve of 1992 actual mortality rates is significantly lower. Values of $R_{2,4,s}$ are thus less than 1.

Table 5.6 Estimated values of $R_{t,4,s}$ used to adjust the expected mortality rates

Year	Scenario 2		Scenario 3	
	Males	Females	Males	Females
1992	0,749	0,795	0,765	0,796
1993	0,925	0,901	0,937	0,898
1994	0,927	0,888	0,938	0,891

5.6.3 Problems

The expected mortality rates (1991 actual rates) involve a number of problems:

- The choice of an expected table of mortality rates is rather subjective. The estimated number of AIDS-related deaths is dependent on the expected rates chosen. Although the 1991 actual mortality rates seem to be an obvious choice, they may not result in the accurate number of AIDS-related deaths.

- Changes in the level of actual mortality curves are generally not uniform across all ages, and the adjustment of expected mortality rates by a constant factor may not be reasonable. For example, if annual fluctuations are due to mortality improvements, these are likely to be less significant at the very old ages, as death cannot be postponed indefinitely. The adjusted expected mortality rates may thus be too low, resulting in an artificial excess of deaths at these ages. However, AIDS-related deaths occur at those ages at which mortality improvements are expected to be relatively uniform, and the problem can thus be largely ignored for the purposes of this dissertation.

5.7 The Estimated Number of Black AIDS-Related Deaths

AIDS-related mortality was estimated as follows for $t = 2, 3, 4$ and $s = 0, 1$:

$$
\begin{aligned}
{}_5q_{t,4,s,x}^{AIDS} &= 0 && \text{for } x = 0, 5, 15, 55, 60, 65, \ldots \\
{}_5q_{t,4,s,x}^{AIDS} &= {}_5q_{t,4,s,x}^{A} - {}_5\tilde{q}_{t,4,s,x}^{E} && \text{for } x = 20, 25, 30, \ldots, 50 \qquad (5.44)
\end{aligned}
$$

Equation 5.44 may result in ${}_5q_{t,4,s,x}^{AIDS} < 0$ for some x. This is likely to be the result of random error, which is assumed to cancel out over the age range 20 to 55.

AIDS-related mortality occurring at ages outside the age range 20 to 55 will not be identified by equation 5.44. It was seen necessary to restrict the estimation of AIDS-related mortality to this range of ages since the expected mortality rates become erratic at older ages.

The number of AIDS-related deaths was initially estimated as shown in equation 5.45:

$$
{}_5AIDS_{t,4,s,x} = {}_5N_{t,4,s,x} \cdot {}_5q_{t,4,s,x}^{AIDS} \qquad (5.45)
$$

Since ${}_5q_{t,4,s,x}^{AIDS}$ may be less than zero, equation 5.45 may yield negative AIDS-related deaths. Estimation of the change in the level of actual mortality rates may

be distorted by changes in the shape of the mortality curve. For example, an improvement in mortality may be underestimated due to increased mortality from AIDS-related causes. AIDS-related deaths may thus be underestimated. In order to correct for this possible underestimation, an iterative procedure was performed:

Let the initial estimates derived in equation 5.45 be denoted ${}_5AIDS_{t,4,s,x}^{(0)}$, and let the corresponding estimate after k iterations be denoted ${}_5AIDS_{t,4,s,x}^{(k)}$. The superscript *(k)* will be used to indicate the value of a variable after k iterations.

The iterative procedure is an extension of the method described in section 5.6.2, and is outlined below:

The ratio of the actual mortality rate from all causes except AIDS to the expected mortality rate after k iterations is given by equation 5.46:

$$ {}_5R_{t,4,s,x}^{(k)} = \frac{{}_5q_{t,4,s,x}^{NonAIDS(k)}}{{}_5q_{t,4,s,x}^{E}} \qquad (5.46) $$

where ${}_5q_{t,4,s,x}^{NonAIDS(0)} = {}_5q_{t,4,s,x}^{A}$

${}_5R_{t,4,s,x}^{(k)}$ was used to recalculate equation 5.40. Equations 5.41, 5.44 and 5.45 were then recalculated (using successive substitutions) after k iterations.

An estimate of mortality from all causes except AIDS was calculated after k+1 iterations as shown in equation 5.47:

$$ {}_5q_{t,4,s,x}^{NonAIDS(k+1)} = {}_5q_{t,4,s,x}^{A} - {}_5q_{t,4,s,x}^{AIDS(k)} \qquad (5.47) $$

This revised estimate is then substituted into equation 5.46 in order to execute the next iterative step. The process is halted once

$$ \sum_{x=20}^{26} {}_5AIDS_{t,4,s,5(x-20)+20}^{(k)} = \sum_{x=20}^{26} {}_5AIDS_{t,4,s,5(x-20)+20}^{(k+1)} $$

to the nearest whole number.

The results are shown in Tables 5.7 and 5.8, and are illustrated in Figure 5.20.

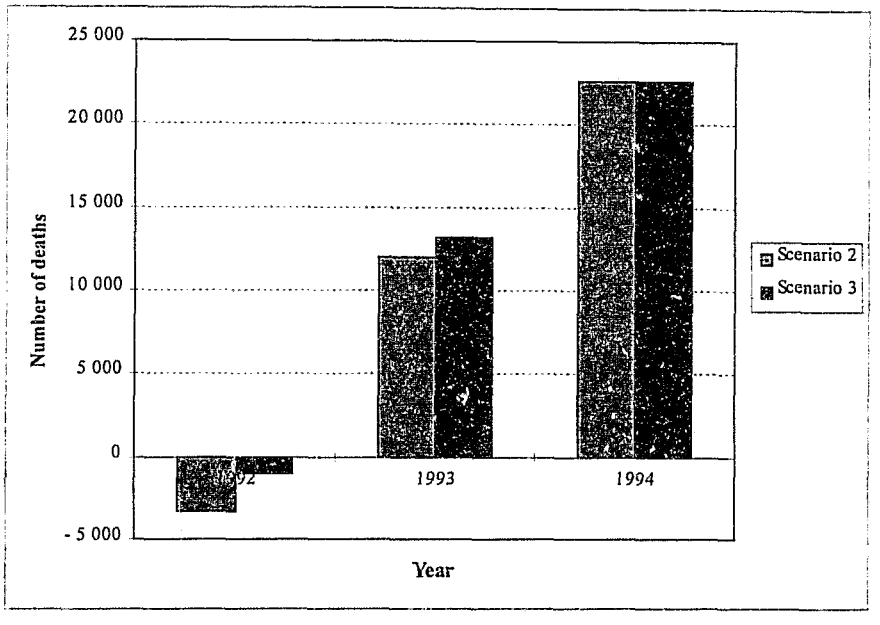


Figure 5.20 Estimated number of black AIDS-related deaths under Scenarios 2 and 3

Table 5.7 Estimated number of black AIDS-related deaths under Scenario 2

Age group	1992		1993		1994	
	Males	Females	Males	Females	Males	Females
20-25	- 1 698	- 130	1 011	714	1 577	1 320
25-30	- 1 278	132	1 134	1 081	2 304	1 902
30-35	- 1 028	311	1 069	1 079	2 221	1 711
35-40	- 728	51	971	612	1 305	1 382
40-45	- 255	631	669	1 096	1 606	1 901
45-50	- 379	500	375	929	1 785	1 569
50-55	425	182	718	558	1 131	894
Total	- 4 941	1 677	5 948	6 070	11 928	10 680

Table 5.8 Estimated number of black AIDS-related deaths under Scenario 3

Age group	1992 Males	1992 Females	1993 Males	1993 Females	1994 Males	1994 Females
20-25	- 1 378	- 495	1 107	716	1 602	1 265
25-30	- 973	192	1 257	1 054	2 299	1 800
30-35	- 742	364	1 208	1 070	2 234	1 645
35-40	- 473	144	1 118	664	1 402	1 359
40-45	- 49	659	832	1 085	1 663	1 816
45-50	- 164	551	553	937	1 814	1 519
50-55	581	313	881	660	1 242	964
Total	- 3 198	2 174	6 955	6 185	12 256	10 369

5.7.1 Discussion

Fewer AIDS-related deaths are estimated under Scenario 2 than under Scenario 3, except for females in 1994.

The 1992 male results are of little use as the totals are not positive values. This may be due to errors in the 1992 data (evident in Table 5.4 and Figures 5.13 to 5.16), as well as distortions involved in the method of estimation, including distortions caused by using year of reporting rather than year of death. The 1992 male results will not be discussed any further.

If the 1992 actual mortality rates had been used as expected mortality rates (rather than the 1991 actual rates), the estimated AIDS-related deaths would be considerably larger in 1993 and 1994. The estimates shown in Tables 5.7 and 5.8 may thus be understated.

The Age Distribution of AIDS-Related Deaths

In order to comment on the reasonability of the age distribution of the estimates in Tables 5.7 and 5.8, some empirical evidence is required. The age and gender distribution of correctly reported AIDS-related deaths in 1992, 1993 and 1994 was obtained from the CSS (1992, 1993 b and unpublished), and is illustrated in Figure 5.21. If the unreported or incorrectly reported AIDS-related deaths are assumed to have the same age structure as the correctly reported deaths, then the results shown in Tables 5.7 and 5.8 should exhibit a similar structure to that shown in Figure 5.21.

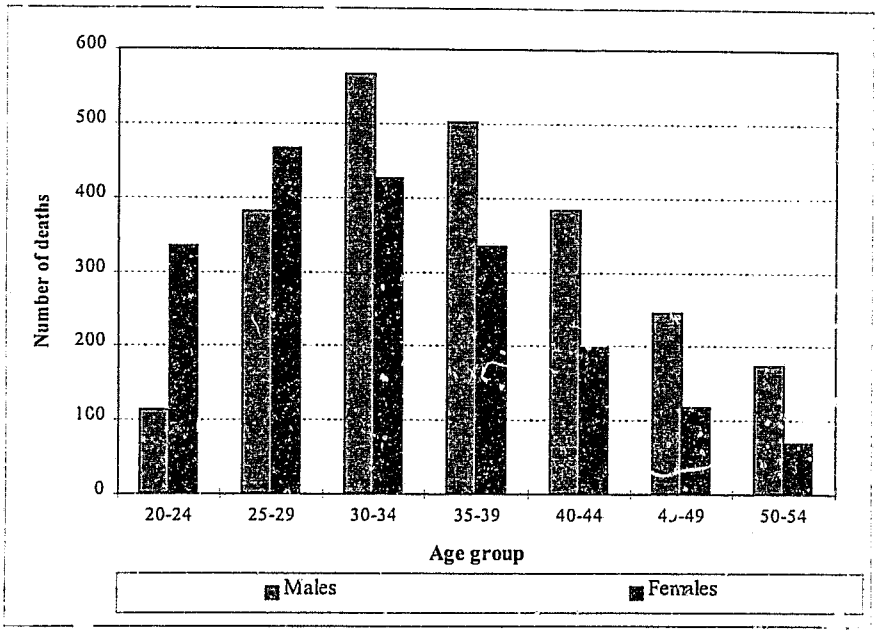


Figure 5.21 The age structure of correctly reported AIDS-related deaths in 1992, 1993 and 1994 combined

Figure 5.21 indicates that the highest number of male reported AIDS-related deaths occurred within the age group 30-35, in comparison to age group 25-30 for females. A gradual decline in AIDS-related deaths occurs with age.

The results in Tables 5.7 and 5.8 do not follow the age distribution illustrated in Figure 5.21. Male deaths peak within the age group 25-30, and female deaths within the age group 40-45 in 1992 and 1993, and 25-30 in 1994. The decline with age is not smooth. For example, a hump occurs between the ages of 40 and 50 for both genders. This is unrealistic, and is due to errors in the underlying data, the random nature of deaths, and further distortions in the method of estimation. The poor quality of the data is likely to be the greatest source of random error.

The distorting effect of these errors is reduced by aggregating data, so total AIDS-related deaths between the ages of 20 and 55 will be used for calibration purposes.

The Gender Distribution of AIDS-Related Deaths

As is evident from Figure 5.21, a greater number of male AIDS-related deaths have been reported than female deaths. This may indicate that a larger number of AIDS-related deaths have occurred amongst males, or may be due to different reporting patterns between the genders. The gender distribution of individuals

who became infected in the mid-eighties would give an idea as to the expected gender distribution of AIDS-related deaths in the mid-nineties. Empirical evidence suggests that the current male to female ratio of HIV-infection is less than 1 (Department of Health, 1996 and Schoub *et al*, 1990). This ratio may however, have been close to 1 in the earlier stages of the epidemic (Bongaarts, 1989: 116).

Total estimated female AIDS-related deaths outweigh those amongst males during 1993 under Scenario 2. The converse is true during 1994, and during 1993 and 1994 under Scenario 3. The ratio of male to female deaths occurring during 1993 is estimated to be 0,98:1 under Scenario 2, and 1,12:1 under Scenario 3. The corresponding ratios for 1994 are estimated to be 1,12:1 and 1,18:1 respectively.

The gender distribution of estimated AIDS-related deaths does not reconcile fully with the available evidence. In particular, it is unlikely that the proportion of male deaths would be increasing over time. However, it is not unreasonable to accept that the ratio of male to female AIDS-related deaths amongst the black population was slightly greater than 1 in the early nineties.

Concluding Remarks

The estimated AIDS-related deaths for 1993 and 1994 are likely to be more reliable than those estimated for 1992. In addition, the underlying death data used in the estimation process are likely to be the most reliable, as it benefits from improvements in reporting delays, and includes deaths occurring in the former TBVC states. Distortions in the age distribution can be eliminated by aggregating the results in each year. Nevertheless, considerable uncertainty surrounds the estimates due to the poor underlying data. The estimates should thus be used with caution.

The 1992 estimates (particularly for males) incorporate considerable random error, and will be ignored for calibration purposes. The ASSA model will be calibrated using the 1993 and 1994 results, aggregated across ages 20 to 55, by gender.

5.8 Summary

Black reported deaths were extracted from the aggregate reported deaths assuming that asians, coloureds and whites died in line with standard mortality tables. Three scenarios were constructed based on the estimated completeness of reporting of deaths amongst these population groups. The SA 84-86 life tables were graduated at the older ages, and used as standard mortality tables.

Estimated black reported deaths were adjusted to allow for underreporting. The Preston and Coale and the Bennett and Horiuchi methods were used, although the

latter was found to be more appropriate. This method indicated that between 49% and 61% of male deaths and between 46% and 57% of female deaths which occurred amongst the black population during 1994, had been reported during the same year.

The black population was estimated based on the 1991 census results, with an adjustment to incorporate deaths occurring in the former TBVC states. The preliminary 1996 census results were compared against the 1991 census results, and inconsistencies discussed.

Actual mortality rates for the black population were calculated for 1991 to 1994 using the estimated death and population data. The 1991 rates were used as "expected" rates up until age 55, and Dorrington's 1991 mortality table thereafter. The expected rates were then compared against those from the three other years. After an adjustment for annual changes in the overall level of the mortality curve, excess mortality between the ages of 20 and 55 was attributed to AIDS.

AIDS-related deaths amongst individuals aged 20-55 were thus estimated for 1992, 1993 and 1994. Those for 1993 and 1994 are believed to be the most accurate, and indicate that between 10 000 and 15 000 AIDS-related deaths occurred during 1993, and between 20 000 and 25 000 during 1994. The male to female ratio is estimated to be slightly greater than 1.

The estimated AIDS-related deaths are analysed further in the following chapter, and will be used to calibrate the ASSA model in Chapter 7.

6

Further Analysis of AIDS-Related Deaths

In addition to the age and gender of the deceased, certain other characteristics are recorded when a death is reported, such as cause of death, occupation, place of residence, place of death and month of death. In this chapter, the estimated AIDS-related deaths are analysed further, with reference to the first three characteristics mentioned.

In order to maintain consistency with the principal analysis performed in Chapter 5, AIDS-related deaths were analysed between the ages of 20 and 55 using quinquennial age groups. The analyses in this chapter were performed under the assumptions of Scenario 2 (defined in section 5.3.1).

6.1 Analysis of AIDS-Related Deaths by Recorded Cause of Death

AIDS patients do not die directly from AIDS, but usually as a result of an opportunistic infection associated with HIV-infection. Medical practitioners are required to complete a medical certificate in respect of the death of a patient. This certificate requires that cause of death be recorded in three ways:

- The final disease or condition resulting in death,
- Any contributory causes, and
- The underlying cause of death.

The death of an AIDS patient should be recorded with AIDS as the underlying cause of death. However, not all AIDS-related deaths are recorded in this way, and vague, incorrect or incomplete causes of death are often recorded.

The South African Law Commission held a workshop entitled *Medical Certificates in Respect of HIV/AIDS Related Deaths* on 7th February 1997, in order to gauge differences in opinion regarding this matter. What came to light was that

certain doctors withheld information if it was believed to be potentially prejudicial to the deceased or the deceased's family.

In order to compile vital statistics, the CSS records only the underlying cause of death. The classification of deaths since 1978 has been done in accordance with the *Ninth Revision of the International Statistical Classification of Diseases, Injuries and Causes of Death* (ICD-9), published by the World Health Organisation (WHO, 1975). As from 1996, such classification has been performed in accordance with the more recent ICD-10 (WHO, 1992). 999 individual causes of death are defined under ICD-9. Natural causes are grouped into 16 broad categories, and unnatural causes into one category. These broad categories are listed below, with the ICD-9 codes in parentheses:

- Infectious and parasitic diseases (001-139)
- Neoplasms (140-208, 210-239)
- Endocrine, nutritional and metabolic diseases and immunity disorders (209, 240-279)
- Diseases of the blood and blood-forming organs (280-289)
- Mental disorders (290-319, 327)
- Diseases of the nervous system and sense organs (320-326, 328-389)
- Diseases of the circulatory system (390-459)
- Diseases of the respiratory system (460-519)
- Diseases of the digestive system (520-579, 609)
- Diseases of the genito-urinary system (580-608, 610-629)
- Complications of pregnancy, childbirth and the puerperium (630-676)
- Diseases of the skin and subcutaneous tissue (680-698, 700-709)
- Diseases of the musculoskeletal system and connective tissue (710-739)
- Congenital anomalies (740-759)
- Certain conditions arising in the perinatal period (760-779)
- Symptoms, signs and ill-defined conditions (699, 780-799)
- Unnatural causes (800-999)

The death reports published by the CSS provide information on deaths grouped by individual causes of death (i.e. by each ICD-9 code from 001 to 999), as well as by the broader categories listed above.

If medical certificates were completed accurately, AIDS would be stated as the underlying cause of all AIDS-related deaths, and the recorded cause of death would be allocated the ICD-9 code 209. Table 6.1 indicates the number of AIDS-related deaths occurring between the ages 20 and 55 which were recorded correctly in 1992, 1993 and 1994 (CSS, 1992; CSS, 1993 b and CSS, unpublished).

It is probable that the remaining AIDS-related deaths were attributed to some other underlying cause. In order to determine the major causes of death of HIV-infected patients, a study was conducted amongst HIV-positive patients who

had attended the HIV clinic of the Johannesburg Hospital and had died during the period July 1991 to June 1995 (Jentsch, 1996).

Table 6.1 The number of recorded deaths amongst individuals aged 20 to 55 with AIDS as the underlying cause of death

Year	Total	Males	Females
1992	680	406	274
1993	1 194	682	512
1994	2 457	1 317	1 140

Source: Central Statistical Service

Of the 82 causes of death noted for the black individuals, 63,4% were infectious and treatable: tuberculosis (26,8%), bacterial infections (14,6%), cryptococcal meningitis (11,0%) and *Pneumocystis carinii* pneumonia (PCP) (11,0%). The common causes of death amongst the white patients were remarkably different: AIDS dementia complex, PCP, disseminated *Mycobacterium avium* complex infection, and disseminated cytomegalovirus infection. Two patterns of mortality emerged from the study: one similar to that of Americans and Europeans (experienced by the white patients), and one similar to patients elsewhere in Africa (experienced by the black patients).

Tuberculosis was observed to be the most prevalent infectious disease and cause of death amongst the black patients in the study. This is corroborated by studies conducted in other African countries (Lucas *et al*, 1993; Nelson *et al*, 1993; Greenberg *et al*, 1995 and Leroy *et al*, 1995), and a strong association between tuberculosis and HIV-infection is indicated. It is thus possible that a substantial number of South African AIDS-related deaths have been attributed to tuberculosis, with ICD-9 codes 010-018.

In addition to tuberculosis, bacterial infections, cryptococcal meningitis and PCP are classified under the broad heading "infectious and parasitic diseases". A number of AIDS-related deaths amongst the black population are thus likely to be found within this category.

A similar analysis was performed using information on deaths amongst homosexual and bisexual men i[illegible] Francisco, who had been recruited between 1978 and 1980 from muni[illegible] transmitted disease clinics for studies of hepatitis B (Hessol *e*[illegible] serum samples were used to identify which deaths were A[illegible] underlying causes of death stated on the

death certificates had been coded according to the ICD-9 system. 1 194 out of 1 292 deaths were found to be AIDS-related, and 109 of these had been incorrectly recorded on the death certificates, as follows:

- Infectious and parasitic diseases (23; 21,1%)
- Neoplasms (26; 23,9%)
- Endocrine, nutritional and metabolic diseases and immunity disorders (22; 20,2%)
- Diseases of the nervous system and sense organs (3; 2,8%)
- Diseases of the respiratory system (5; 4,6%)
- Diseases of the digestive system (6; 5,5%)
- Unnatural causes (24; 22,0%)

The common causes of death of white San Franciscan males are not necessarily likely causes of death of black South African AIDS patients. The importance of "infectious and parasitic diseases" as a recorded cause of death amongst AIDS patients is nevertheless emphasised. Neoplasms (particularly Kaposi's sarcoma) have been found to be closely associated with white, rather than black, South African AIDS patients (Jentsch, 1996), and relatively few black AIDS-related deaths may have been attributed to this category. Although unnatural causes constitute 22% of the incorrectly recorded causes of death in the above study, 15 of the 24 were due to suicide. This cause of death has been found to be relatively uncommon amongst blacks (Foster and Keen, 1988; Mayekiso, 1995), and unnatural causes are thus assumed to constitute a negligible proportion of the incorrectly recorded AIDS-related deaths amongst the black population of South Africa.

Ill-defined causes of death are often used to measure the quality of the reporting of deaths. A number of papers have drawn attention to the large proportion of South African deaths which are classified in this way (Botha and Bradshaw, 1935; Kielkowski, Steinberg and Barron, 1989; Bradshaw, Dorrington and Sitas, 1992; and Bah and Kleinschmidt, 1997). 23,1%, 21,3% and 16,7% of recorded deaths were attributed to ill-defined causes in 1992, 1993 and 1994 respectively.

van der Merwe, Yach and Metcalf (1991) conducted a study in order to identify the reasons for the high proportion of ill-defined deaths in Port Elizabeth during the year ending 30th June 1989. The majority of ill-defined deaths had been certified by someone other than a medical practitioner (e.g. when cause of death is under investigation, and is certified as being due to "natural causes" by a police officer). There is no reason to suggest that these deaths are more likely to be AIDS-related. The causes of death of 26% of the hospital deaths in the study were not accurately recorded on the medical certificate, and would be coded as "ill-defined" by CSS nosologists. As discussed, such inadequate recording may have been deliberate in the case of HIV-positive patients. It is thus possible that a proportion of AIDS-related deaths have been attributed to "ill-defined causes" in the past.

The evidence suggests that the following recorded causes of death conceal the majority of AIDS-related deaths amongst the black population of South Africa (with abbreviations in brackets):

1. Tuberculosis (TB)
2. Other infectious and parasitic diseases (i.e. excluding tuberculosis) (INF)
3. Neoplasms (NEO)
4. Endocrine, nutritional and metabolic diseases and immunity disorders, excluding the correctly coded AIDS-related deaths (END)
5. Symptoms, signs and ill-defined conditions (ILL)

The variable m will be used to denote cause of death, and assumes the values 1 to 5 as defined above. Although the data allow for analysis by each individual cause of death, this was seen to be spuriously accurate due to misreporting of cause of death. The pattern of misreporting is unlikely to remain constant over time, thus introducing further distortions. The analysis (described in the remainder of section 6.1) was thus performed using broad categories of cause of death.

The 1991, 1992, 1993 and 1994 death reports published by the CSS (1991 a, 1992, 1993 b and unpublished) do not provide information on the population group of the deceased. A method was derived in section 5.3.1 to extract the black reported deaths from the aggregate reported deaths, by assuming that the actual pattern of non-black deaths (i.e. deaths amongst asians, coloureds and whites) was equal to an expected pattern based on standard tables of mortality.

Standard tables of mortality do not exist by cause, and it was thus impossible to extract the black reported deaths from the aggregate reported deaths for each individual cause m. Deaths occurring within all population groups were therefore analysed.

The notation defined in section 5.2 will be used. A superscript m (m = 1, 2, 3, 4, 5 as listed above) will be used to denote the cause of death in operation. For example,

$$_{y-x}D^{m}_{t,i,s,x}$$

denotes the number of deaths which occurred as a result of cause m during year t amongst individuals in population group i, with gender s and aged between x and y.

6.1.1 Adjustment for Underreporting

It is assumed that 90% of non-black deaths were reported in each age group under Scenario 2.

The Bennett and Horiuchi method (section 5.3.3) was chosen to estimate the completeness of reporting of black deaths. This method relies upon the assumption of a uniform completeness of reporting across the adult ages.

The completeness of reporting of deaths in each population group was thus assumed to be uniform across all adult ages. However, this does not mean that the completeness of reporting is uniform when the population groups are combined. A change in the ratio of non-black to black deaths would distort the completeness of reporting of deaths. This is likely to occur, as non-blacks (particularly asians and whites) die at older ages than blacks, thus increasing the completeness of reporting at these ages. Separate estimates were thus derived for each quinquennial age group between 20 and 55.

The completeness of reporting of deaths occurring in year t (t = 1, 2, 3, 4) amongst individuals from all population groups, with gender s (s = 0, 1), and aged between x and x+5 (x = 20, 25, 30, ..., 50) was calculated as shown in equation 6.1:

$$ {}_5\overline{C}_{t,s,x} = \frac{C_{t,4,s} \cdot {}_5D_{t,4,s,x} + 0{,}9 \cdot \sum_{i=1}^{3} {}_5N_{t,i,s,x} \cdot {}_5q_{i,s,x}{}^{s}}{{}_5\overline{D}_{t,s,x}} \qquad (6.1) $$

$C_{t,4,s}$ is the Bennett and Horiuchi estimate of the completeness of reporting of deaths occurring in year t amongst individuals in the black population with gender s, as calculated in section 5.3.4.

The estimates are plotted against the midpoint of each quinquennial age group in Figure 6.1.

The following is noted:

- As expected, the completeness of reporting increases at the older ages for both genders as the ratio of non-black deaths to black deaths increases.
- The completeness of reporting of male deaths declines from age 20 to age 30, before rising again. The proportion of non-black deaths relative to black deaths is greater during the early twenties due to the so-called "accident hump" which has an effect at these ages. This decline is not as significant for females, who typically experience a less significant "accident hump".
- The distortions mentioned in sections 5.3.1 and 5.3.2 are relevant to this analysis, and should be borne in mind when interpreting the results.

In order to adjust the reported deaths from each cause m for underreporting, it was necessary to assume that the completeness of reporting was uniform across all causes of death. This is an unrealistic assumption. Causes inducing the death of individuals with a higher standard of living are more likely to be reported than

those (often treatable) causes which induce the death of individuals living in poorer conditions. A greater proportion of deaths from unnatural causes may be reported than from natural causes, as police investigation is often required. The assumption thus overstates mortality from unnatural causes and those common within the higher socio-economic classes, and understates mortality from those causes which induce death within the lower socio-economic classes.

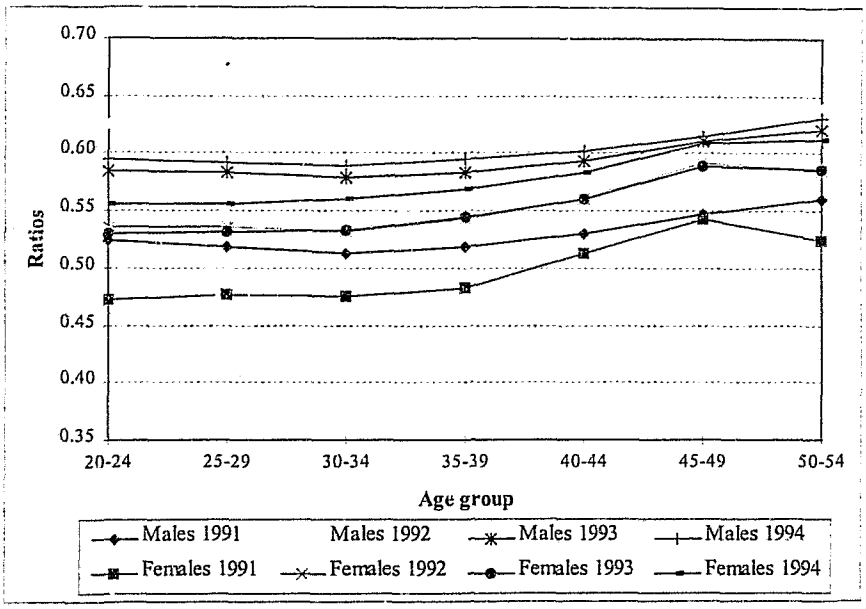


Figure 6.1 Estimated completeness of reporting of deaths in all population groups

The actual number of deaths from cause m occurring in year t amongst individuals with gender s and aged between x and x+5 was estimated as shown in equation 6.2:

$$ {}_5\overline{D}^{m}_{t,s,x} = \frac{{}_5\overline{D}^{m\ R}_{t,s,x}}{{}_5\overline{C}_{t,s,x}} \qquad (6.2) $$

Values of ${}_5\overline{D}^{m\ R}_{t,s,x}$ were obtained from the CSS death reports.

6.1.2 Analysis of Cause-Specific Mortality Rates

The five different causes of death under consideration constitute part of a multiple

decrement table.

The actual average probability for year t that an individual in population group i, with gender s and aged between x and $x+5$ will die within one year as a result of cause m was estimated as shown in equation 6.3:

$$ {}_5\overline{q}^{m\ A}_{t,s,x} = \frac{{}_5\overline{D}^{m}_{t,s,x}}{{}_5\overline{N}_{t,s,x} + 1/2.{}_5\overline{D}_{t,s,x}} \qquad (6.3) $$

This equation is of the form of equation 5.35. A discussion of multiple decrement tables and the derivation of equation 6.3 may be found in Neill (1977: 291-320).

As discussed in section 5.1, the AIDS epidemic is expected to change the shape of the mortality curve. In particular, mortality is expected to worsen significantly across the ages 20 to 55. The mortality rates for certain causes of death may exhibit such a pattern, and may thus conceal a number of AIDS-related deaths. Caution must be exercised when attributing worsening mortality to AIDS, since the prevalence of certain diseases may be increasing for reasons unrelated to AIDS. For example, an increase in mortality from circulatory causes and neoplasms may be the result of socio-economic development, as they are related to the modern lifestyle (Bah and Kleinschmidt, 1997).

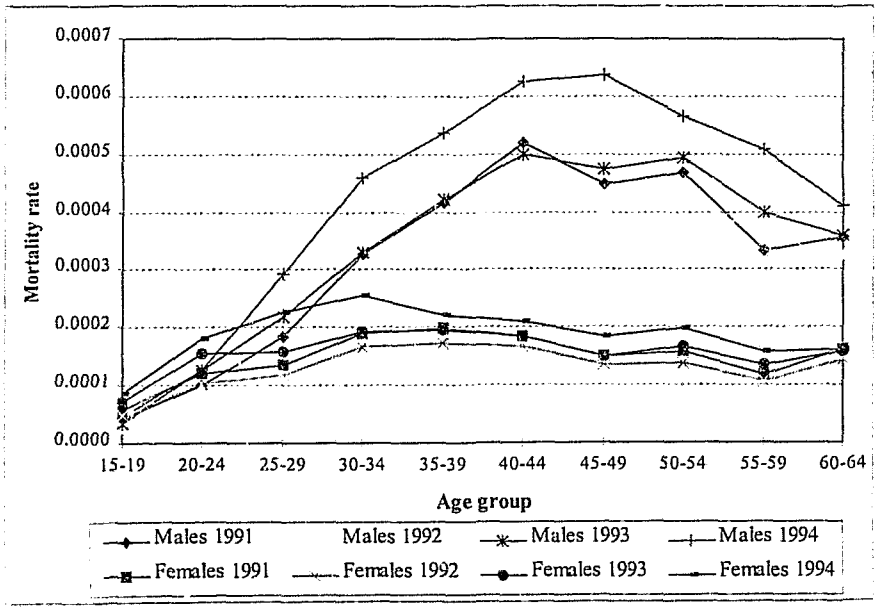


Figure 6.2 Estimated actual rates of mortality from tuberculosis (TB)

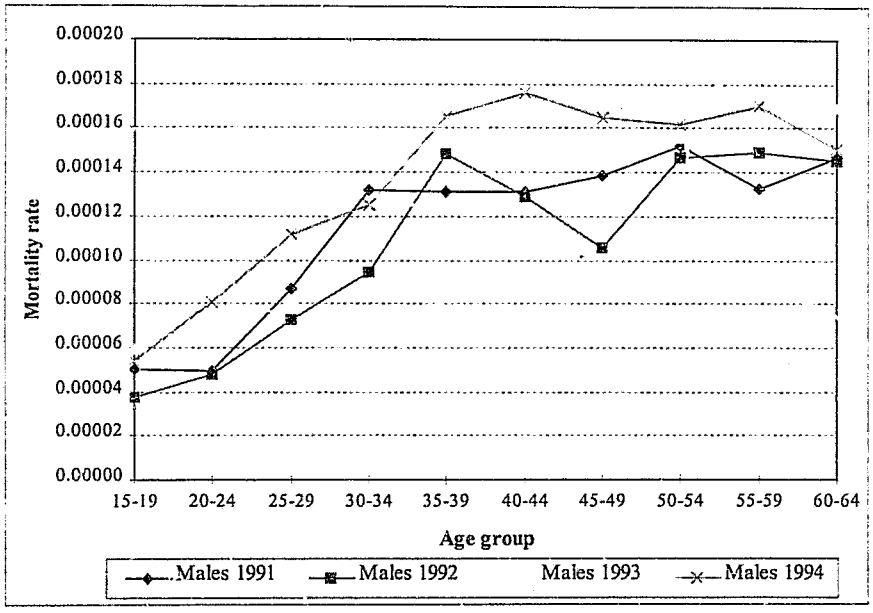


Figure 6.3 Estimated actual rates of mortality from other infectious and parasitic diseases (INF), for males

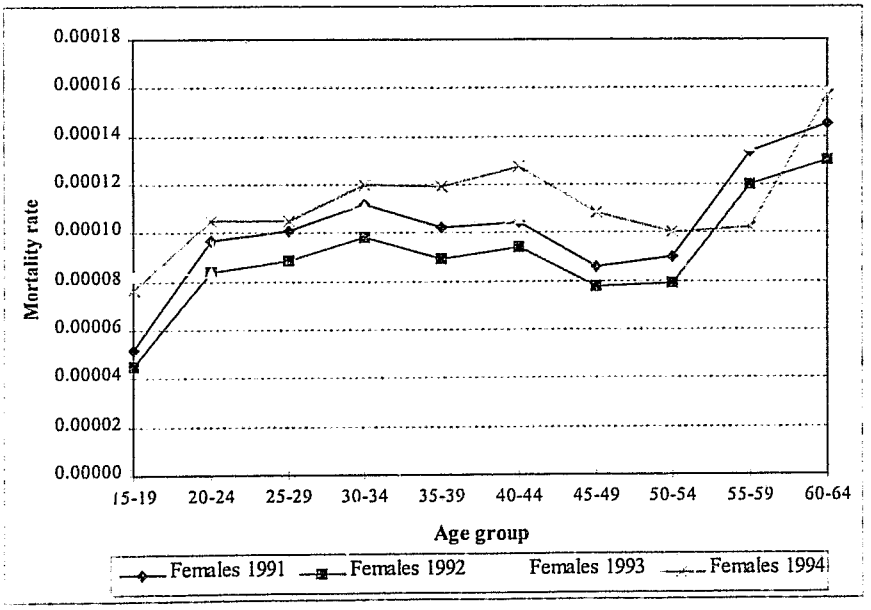


Figure 6.4 Estimated actual rates of mortality from other infectious and parasitic diseases (INF), for females

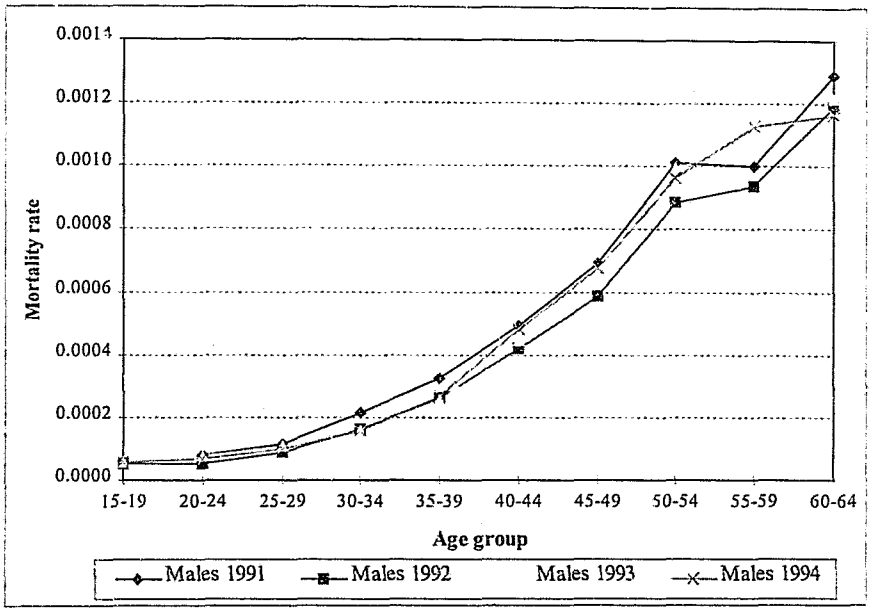


Figure 6.5 Estimated actual rates of mortality from neoplasms (NEO), for males

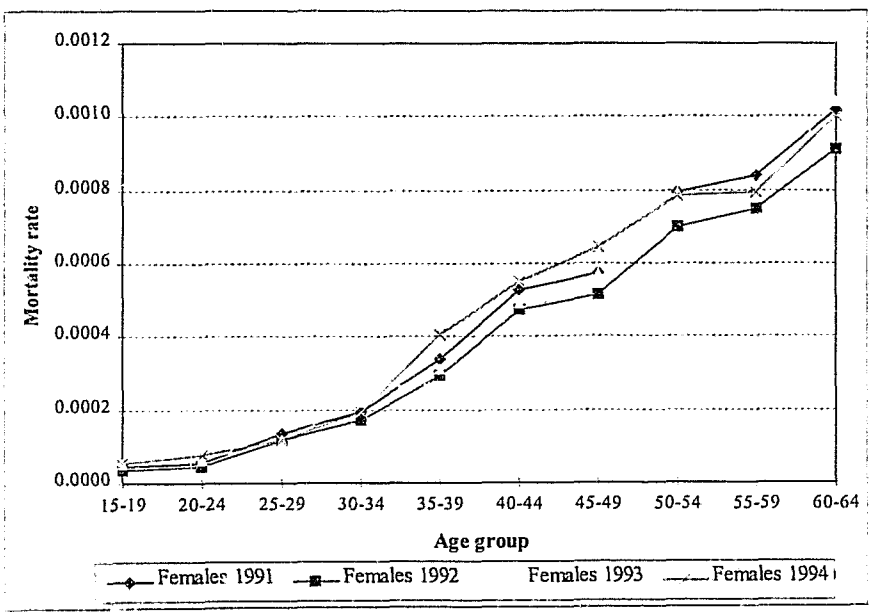


Figure 6.6 Estimated actual rates of mortality from neoplasms (NEO), for females

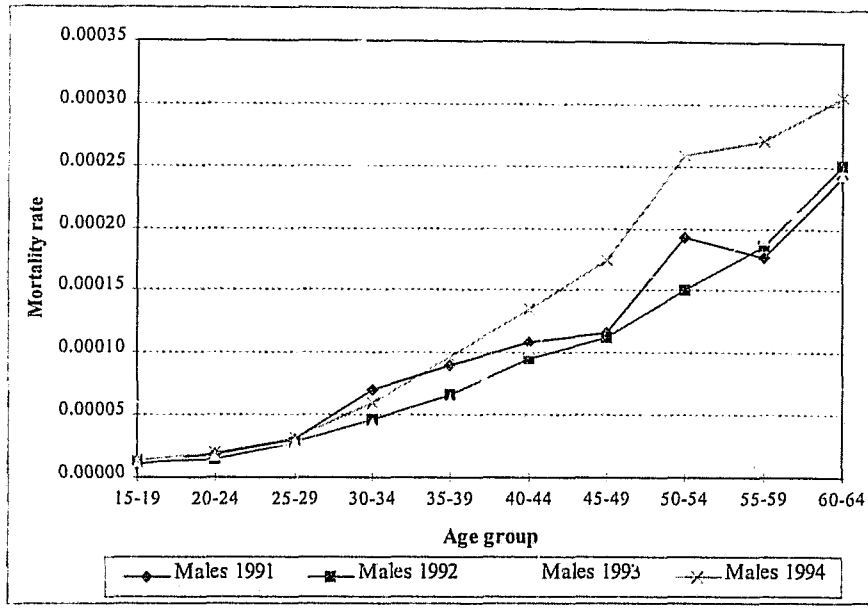


Figure 6.7 Estimated actual rates of mortality from endocrine, nutritional and metabolic diseases and immunity disorders (END), for males

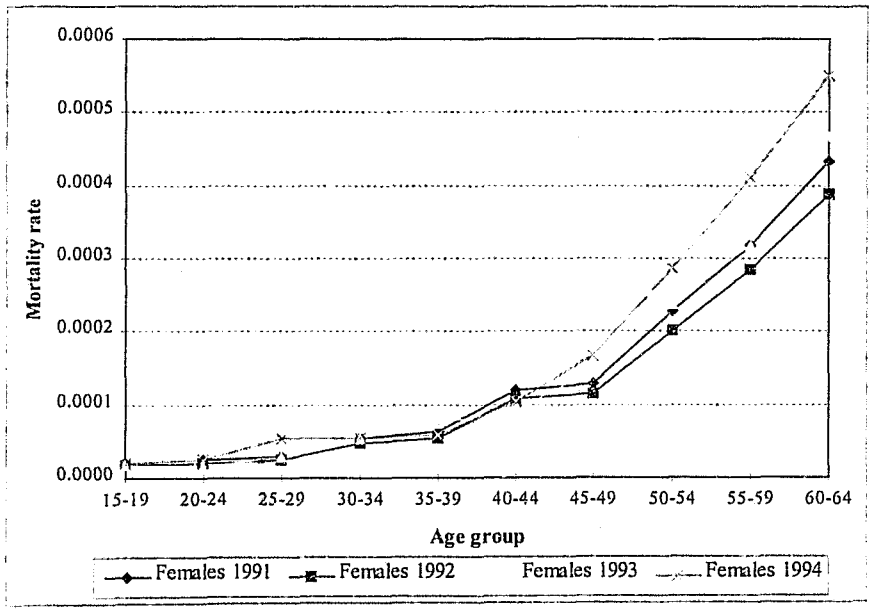


Figure 6.8 Estimated actual rates of mortality from endocrine, nutritional and metabolic diseases and immunity disorders (END), for females

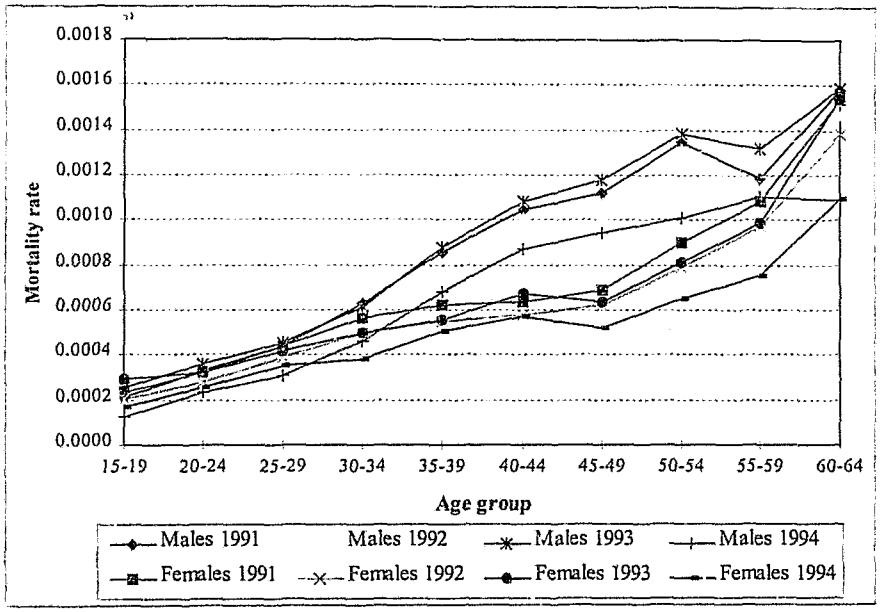


Figure 6.9 Estimated actual rates of mortality from symptoms, signs and ill-defined conditions (ILL)

The mortality curves for the causes of death denoted $m = 1, 2, ..., 5$ are shown in Figures 6.2 to 6.9.

The following observations are made:

- The 1992 mortality rates are unusually low across all causes of death. This is consistent with the 1992 population rates calculated in section 5.5, and is probably due to errors in the underlying data.
- The incidence of mortality due to tuberculosis decreased between 1991 and 1992, and then increased significantly. The largest increase occurs at ages 40 to 50 for males, and 20 to 30 for females.
- Mortality due to other infectious and parasitic diseases appears to have increased since 1992, although rather erratically.
- Mortality as a result of neoplasms has increased visibly for females in 1994 across the ages 30 to 50. Male mortality does not show as obvious a trend.
- Mortality due to endocrine, nutritional and metabolic diseases and immunity disorders increased in 1993 and 1994, particularly at ages above 40.
- Mortality due to symptoms, signs and ill-defined conditions decreased between 1991 and 1992, increased from 1992 to 1993, and then dropped substantially in 1994. This indicates an improvement in the quality of reporting between 1993 and 1994. This reported cause of death will not be analysed further since changes in reporting practices affect the results.

- The better classification of deaths in 1994 may explain the increased mortality observed for specific causes during that year.

6.1.3 Results

The "excess" deaths resulting from an increase in mortality for reported cause m between 1991 and year t (denoted as ${}_5\overline{INCR}^{m}_{t,s,x}$) were estimated for t = 3, 4 and between the ages of 20 and 55 as shown in equation 6.4:

$$ {}_5\overline{INCR}^{m}_{t,s,x} = ({}_5\overline{q}^{m\ A}_{t,s,x} - {}_5\overline{q}^{m\ A}_{1,s,x})\,{}_5\overline{N}_{t,s,x} \qquad \text{for } x = 20, 25, 30, \ldots, 50 $$

$$ {}_5\overline{INCR}^{m}_{t,s,x} = 0 \qquad \text{elsewhere} \qquad (6.4) $$

No adjustment has been made for annual changes in the level of mortality from cause m (as was done in section 5.6.2) as it would make little difference to the results. The 1991 actual rates were used as expected rates even at the older ages where they are more erratic. This was done as no standard tables exist by cause with which to smooth the actual rates.

These "excess" deaths may provide some clue as to the cause of death which was recorded in respect of AIDS patients. Tables 6.2 and 6.3 indicate the proportion of the AIDS-related deaths estimated in Chapter 5 which can be explained by these "excess" deaths under the assumptions of Scenario 2.

Table 6.2 Breakdown of estimated adult, male AIDS-related deaths by recorded cause of death

	1993	1994
AIDS-related deaths (Chapter 5)	**5 948**	**11 928**
Recorded deaths	682	1 317
Percentage recorded	*11,5%*	*11,0%*
TB	197	1 537
INF	41	326
NEO	-465	-422
END	-107	295
Percentage incorrectly recorded	*-5,6%*	*14,6%*
Percentage "explained"	***5,8%***	***25,6%***

Table 6.3 Breakdown of estimated adult, female AIDS-related deaths by recorded cause of death

	1993	1994
AIDS-related deaths (Chapter 5)	**6 070**	**10 680**
Recorded deaths	512	1 140
Percentage recorded	*8,4%*	*10,7%*
TB	131	673
INF	42	183
NEO	-78	313
END	16	201
Percentage incorrectly recorded	*1,8%*	*12,8%*
Percentage "explained"	***10,3%***	***23,5%***

The 1994 results presented in Tables 6.2 and 6.3 are assumed to be more accurate than the 1993 results for the following reasons:

- The quality of recording of cause of death appears to have improved in 1994, as evidenced by the reduction in the number of ill-defined causes.
- The percentages in the tables are greater than 0%.

6.1.4 Discussion

All conclusions are drawn from the 1994 results.

Approximately 11% of estimated AIDS-related deaths were reported to the authorities. This is consistent with the 10% reporting rate found in other African countries (Schoub *et al*, 1990).

Of those incorrectly classified AIDS-related deaths which can be explained, the largest proportion appear to have been attributed to TB. This is not unexpected, as a strong relationship between HIV infection and TB has been documented in other African countries.

A relatively large proportion of incorrectly classified female AIDS-related deaths appear to have been attributed to neoplasms. Cancer of the cervix has been found to be the most common cancer in black females (Sitas, Terblanche and Madhoo,

1996). However, no association has been found between cervical cancer and HIV infection, both in South Africa (Sitas *et al*, 1997) and Tanzania (ter Meulen *et al*, 1992). The findings may thus be the result of random error, and further investigation regarding the role of neoplasms in the death of black female AIDS patients is required.

74% and 77% of the AIDS-related deaths occurring during 1994 amongst males and females respectively cannot be explained using the above analysis. This may be explained in a number of ways:

- The 1994 estimates of AIDS-related deaths are too high, especially for females.
- The four causes (TB, INF, NEO, and END) do not account for all the incorrectly recorded AIDS-related deaths. especially those in respect of females. This suggests that the currently unexplained AIDS-related deaths were attributed to the other causes of death listed in section 6.1.
- As discussed in section 6.1.2, the assumption of uniform underreporting across each of the causes is likely to understate mortality attributable to common causes of death amongst the lower socio-economic classes. A large proportion of the black population, particularly the female population, would be classified into one of the lower socio-economic classes. The origin of the unexplained deaths, as well as the difference between males and females can thus be justified.
- Equation 6.4 does not produce realistic estimates of the number of AIDS-related deaths attributed to the four causes. For example, although AIDS may increase mortality from cause m at certain ages, other factors (such as immunisation, improvement in medical services, etc.) may counteract this effect. The shape of the mortality curve from cause m may thus remain constant.

The last two of the abovementioned points are believed to be particularly relevant to the analysis.

6.2 Analysis of AIDS-Related Deaths by Occupation

It is well known that an individual's mortality is affected by his/her nutrition and surrounding environment. These two factors can be controlled depending on the individual's level of affluence, which is largely dependent on occupation. Occupation may also be an indicator of behaviour, since highly paid jobs are generally held by the more educated individuals who tend to exhibit behaviour which lowers their mortality risk (e.g. regular exercise, moderate intake of alcohol, moderate smoking habits, etc.).

Individual behaviour is the most critical factor determining the risk of HIV transmission. In the light of the association between occupation and individual

behaviour, different occupation groups are expected to display different levels of HIV prevalence, and thus, different rates of AIDS-related mortality. Since HIV is spread by sexual contact, those occupations which present opportunities for such contact are likely to experience higher levels of prevalence.

African HIV prevalence appears to be higher amongst economically active individuals as opposed to economically inactive individuals. Economically active individuals tend to live in urban areas, where higher prevalence rates have been extensively noted (Barnett and Blaikie, 1992 and Over, 1990:2). According to a WHO study, rural HIV prevalence rates are between 10% and 55% of urban prevalence rates (WHO, 1989:15).

Economically active individuals enjoy higher socio-economic status, which is also associated with higher prevalence rates (Ijsselmuiden, 1992). In a Rwandan sample, approximately one third of HIV-infected individuals had completed secondary or higher education. Similar findings have been noted in Zambia and Uganda (Barnett and Blaikie, 1992:26, 27, 47). Professionals and other "elite" classes are thus expected to experience higher mortality rates.

HIV prevalence amongst commercial sex workers is likely to be extremely high, by the very nature of their activity. Prevalence rates of 80% (Nairobi, Kenya), 55% (Abidjan, Côte d'Ivoire) and 55% (Djibouti) have been documented (UNAIDS, 1996), and rates as high as 90% in Kigali, Rwanda (Fleming, 1993). Prevalence amongst Tanzanian prostitutes was found to be 39,0% in 1989, compared to a prevalence of 9,8% amongst Tanzanian antenatal clinic attendees during the same year (Doyle, 1991).

The movement of individuals is associated with a higher risk of HIV infection, as it presents opportunities for new sexual contacts. This has been confirmed in the case of long-distance truckers (Fleming, 1993). 26,7% of 952 truck drivers on the Mombasa-Nairobi highway were found to be HIV-positive during 1991 (Ijsselmuiden, 1992). High prevalence levels have been documented along the Nairobi-Kampala-Kigali highway (Brown, 1990). Increasing levels of HIV prevalence have been found on major truck routes along the northern shores of Lake Victoria (Okware, 1987), suggestive of interaction with infected truckers.

In Southern Africa, large-scale migration of individuals occurs for employment reasons. Large numbers of males leave their homes and families in order to work on the gold mines. These males are housed in hostels, and tend to suffer from loneliness and social dislocation (Whiteside, 1993). The reliance upon prostitutes is believed to be considerable (Hambridge, 1990). The HIV prevalence amongst South African miners was estimated to be between 10% and 20% in 1995, with substantial variation between mines (Foster, 1996). This is somewhat higher than the estimated prevalence amongst the adult population of 7,8% (Doyle, 1996).

Military personnel are extremely susceptible to HIV infection for a number of

reasons (Fleming, 1993; Whiteside, 1993; Miller and Yeager, 1996 and *AIDS Analysis Africa (Southern Africa Edition)*, 1996). They comprise the sexually active age group 15 to 24, are away from home for extended periods and are exposed to commercial sex workers. Feelings of invincibility and a tendency toward risk-taking are believed to be common. The HIV prevalence in the Angolan government and UNITA troops is estimated to be 50% compared to a prevalence of 10% in the general population. Similarly, Zimbabwean soldiers display an HIV prevalence of 50% compared to 10% to 20% in the population (*AIDS Analysis Africa (Southern Africa Edition)*, 1996).

In the compilation of vital statistics, the CSS defines occupation as "the activity or collection of activities which constitute the tasks or type of work done by a person during his day's work". Thus, officially unemployed people who conduct some sort of work during the day (e.g. informal traders) are assumed to be economically active. Occupations are classified into ten major occupation groups in accordance with the *Standard Classification of Occupations* published by the CSS (1986). These occupation groups are listed below:

1. Professional, semi-professional and technical occupations
2. Managerial, executive and administrative occupations
3. Clerical and sales occupations
4. Transport, delivery and communications occupations
5. Service occupations
6. Farming and related occupations
7. Artisan, apprentice and related occupations
8. Production supervisor, miner, quarry and related worker
9. Occupation unspecified and not elsewhere classified
10. Not economically active

Military personnel are grouped with other "protective service occupations" under occupation group 5. Similarly, escorts (and perhaps prostitutes) are grouped under occupation group 5 with other "personal and related services". Occupation group 9 is reserved for those occupations which cannot be classified in groups 1 to 8, those occupations which are not sufficiently described on the death certificate, and those individuals who are actively seeking work but cannot find anything suitable. Prostitutes may thus be classified under this occupation group. Occupation group 10 is reserved for individuals who are not actively seeking work (including children and the elderly).

As mentioned in section 6.1, the 1991, 1992, 1993 and 1994 death reports published by the CSS (1991 a, 1992, 1993 b and unpublished) do not provide information on the population group of the deceased. In contrast to the method applied in section 5.3.1, it was not possible to extract the black reported deaths from the aggregate reported deaths for each occupation group, as standard tables of mortality do not exist by occupation group. Thus, as in section 6.1, deaths from all population groups were analysed.

The notation defined in section 5.2 will be used. A superscript j (where j = 1, 2, 3, ..., 10) will be used to denote occupation group, as listed above. For example,

$$_{y-x}{}^{j}N_{t,i,s,x}$$

denotes the number of individuals in population group i, occupation group j, with gender s and aged between x and y on 30th June of year t.

6.2.1 Adjustment for Underreporting

The same estimates of the completeness of reporting ($_5\overline{C}_{t,s,x}$) were used as described in section 6.1.1.

It was necessary to assume that the completeness of reporting was the same across all occupation groups. This is unrealistic, as a marked difference is expected between economically active and economically inactive individuals. Economically active individuals generally belong to a retirement fund, and a death is likely to be reported by the beneficiaries in order to claim benefits. Moreover, differences in the completeness of reporting may exist between the various occupation groups constituting the economically active population. The assumption is likely to overstate the mortality experience of the economically active population and understate that of the economically inactive population. The level of overstatement varies between the individual economically active occupation groups depending on the relative differences in the completeness of reporting.

The actual number of deaths occurring in year t amongst individuals in occupation group j, with gender s and aged between x and x+5 was estimated as follows:

$$_5^j\overline{D}_{t,s,x} = \frac{_5^j\overline{D}_{t,s,x}{}^{R}}{_5\overline{C}_{t,s,x}} \qquad (6.5)$$

Values of $_5^j\overline{D}_{t,s,x}{}^{R}$ were obtained from the CSS death reports. Equation 6.5 is analogous to equation 6.2.

6.2.2 Estimation of the Population by Occupation Group

In order to adhere to the principle of correspondence, estimates of the population by occupation group are required. Ideally, these estimates should be subdivided by gender and age.

The published 1991 census results do not present a breakdown by occupation group, gender and age simultaneously. A breakdown by occupation group and age

is, however, available (CSS, 1991 d).

The population was classified into the broad age bands: (0, 20), (20, 25), (25, 35), (35, 55), (55, 65) and (65, ω). These estimates were broken down further into quinquennial age bands as follows:

Let ${}^{j}_{b-a}\overline{\overline{N}}_{1,a}$ denote the 1991 census estimate of the population within occupation group j and age group (a, b) (where a and b are the upper and lower limits of the age bands listed above).

${}_{5}\overline{\overline{N}}_{1,x}$ denotes the 1991 estimate of the total population of South Africa within the age group $(x, x+5)$. This estimate was obtained by summing the 1991 census estimates of the asian, coloured, white and black populations, including an estimate of the population living in the former TBVC states in 1991. The latter was described in section 5.4.3.

$$\text{Then, } {}^{j}_{5}\overline{\overline{N}}_{1,x} = {}^{j}_{b-a}\overline{\overline{N}}_{1,a} \cdot \frac{{}_{5}\overline{\overline{N}}_{1,x}}{{}_{b-a}\overline{\overline{N}}_{1,a}} \qquad (6.6)$$

where $x \geq a$ and $y \leq b$.

The proportion of the 1991 population aged between x and $x+5$ which was in occupation group j (denoted ${}^{j}_{5}p_{x}$) was calculated as shown in equation 6.7:

$${}^{j}_{5}p_{x} = \frac{{}^{j}_{5}\overline{\overline{N}}_{1,x}}{{}_{5}\overline{\overline{N}}_{1,x}} \qquad (6.7)$$

These proportions were assumed to remain constant over time (which may be unrealistic) and were used to calculate the size of the population in occupation group j and aged between x and $x+5$ at the midpoint of year t (t = 2, 3, 4) as follows:

$${}^{j}_{5}\overline{\overline{N}}_{t,x} = {}^{j}_{5}p_{x} \cdot {}_{5}\overline{\overline{N}}_{t,x} \qquad (6.8)$$

where ${}_{5}\overline{\overline{N}}_{t,x}$ is made up of the estimates of the black population derived in section 5.4, and the estimates of the asian, coloured and white populations derived in section 5.3.1.

6.2.3 Analysis of Mortality Rates by Occupation

The actual average probability for year t that an individual in occupation group j

and aged between x and x+5 will die within one year, was calculated as shown in equation 6.9:

$$ {}_5\hat{q}_{t,x} = \frac{{}_5\bar{D}_{t,x}}{{}_5\bar{N}_{t,x} + 1/2.{}_5\bar{D}_{t,x}} \qquad (6.9) $$

The derivation of equation 6.9 is analogous to equation 5.35.

Changes in the shape of the mortality curve since 1991 may be indicative of AIDS-related deaths. However, mortality may increase due to a change in the composition of individuals in that occupation group. This is particularly relevant, since the four population groups combined in this analysis are in fact heterogeneous with respect to mortality. As a larger proportion of blacks are entering the workplace to fill positions previously for whites, the mortality experience may change due to differing cultural and expenditure patterns. The results of this analysis are thus subject to considerable uncertainty, and may be meaningless if the composition is found have altered at a rapid rate.

The mortality curves for each occupation group are illustrated in Figures 6.10 to 6.19. Caution must be exercised when interpreting these graphs, since the scales of the y-axes differ.

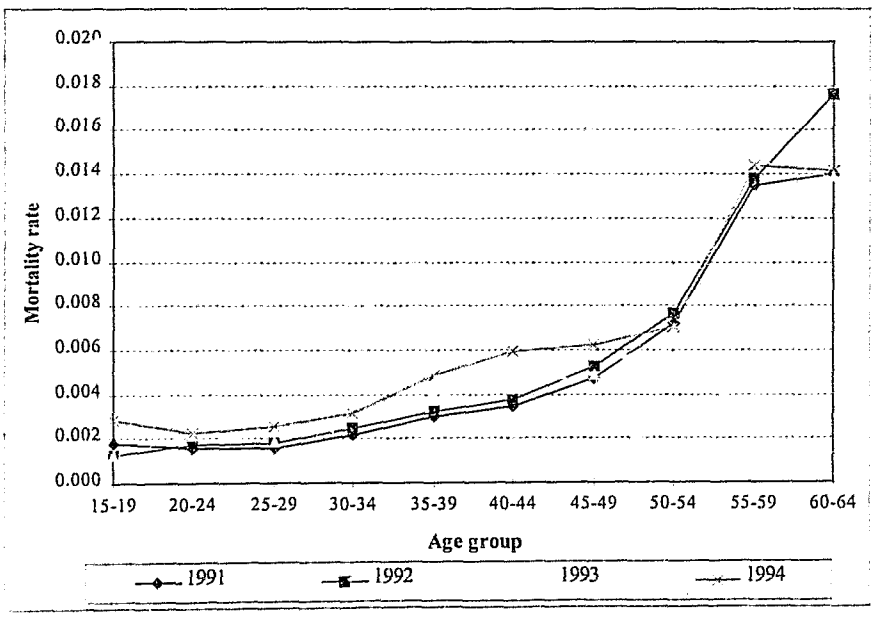


Figure 6.10 Estimated actual rates of mortality for: professional, semi-professional and technical occupations

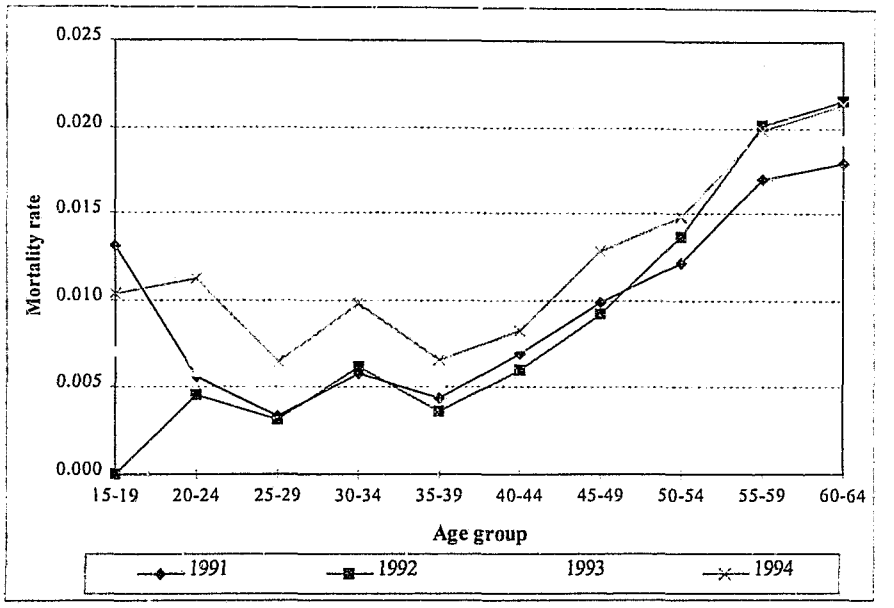


Figure 6.11 Estimated actual rates of mortality for: management, executive and administrative occupations

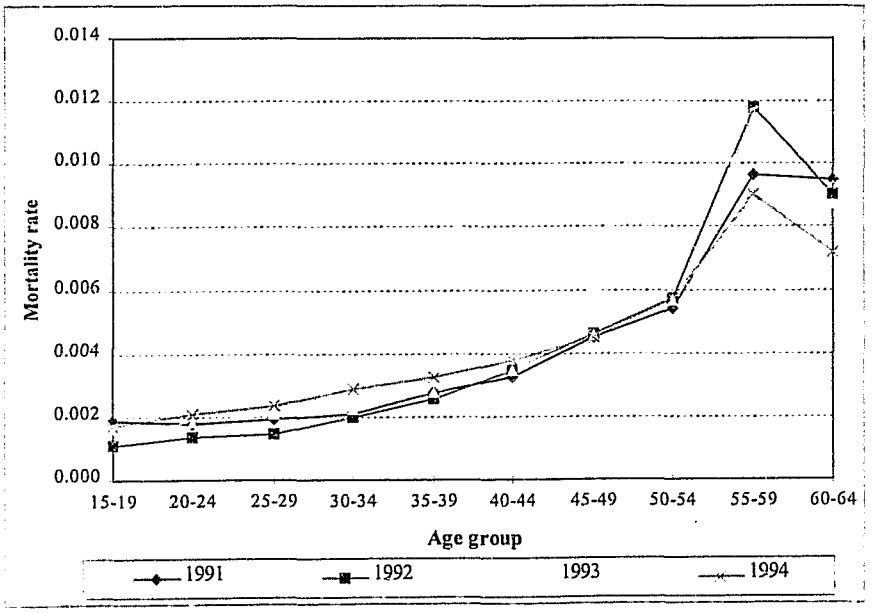


Figure 6.12 Estimated actual rates of mortality for: clerical and sales occupations

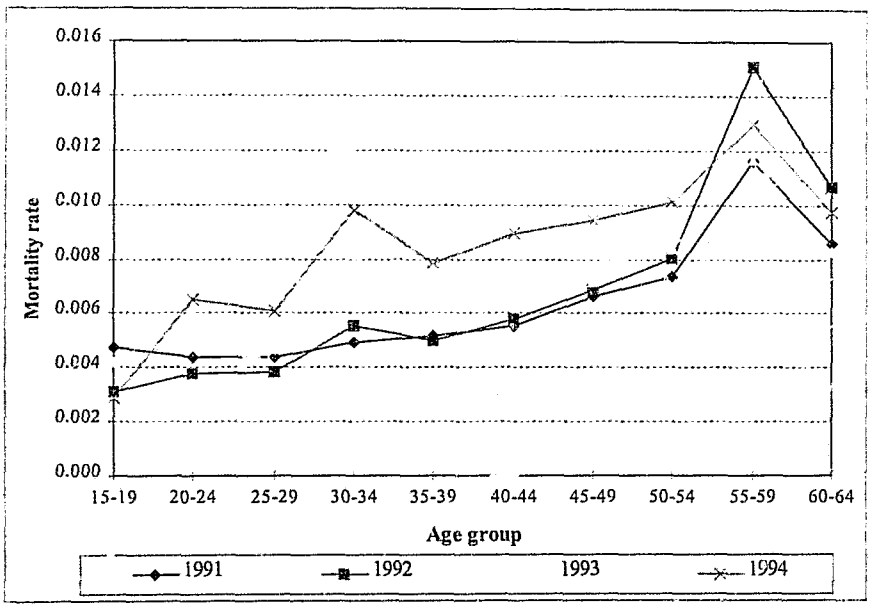


Figure 6.13 Estimated actual rates of mortality for: transport, delivery and communications occupations

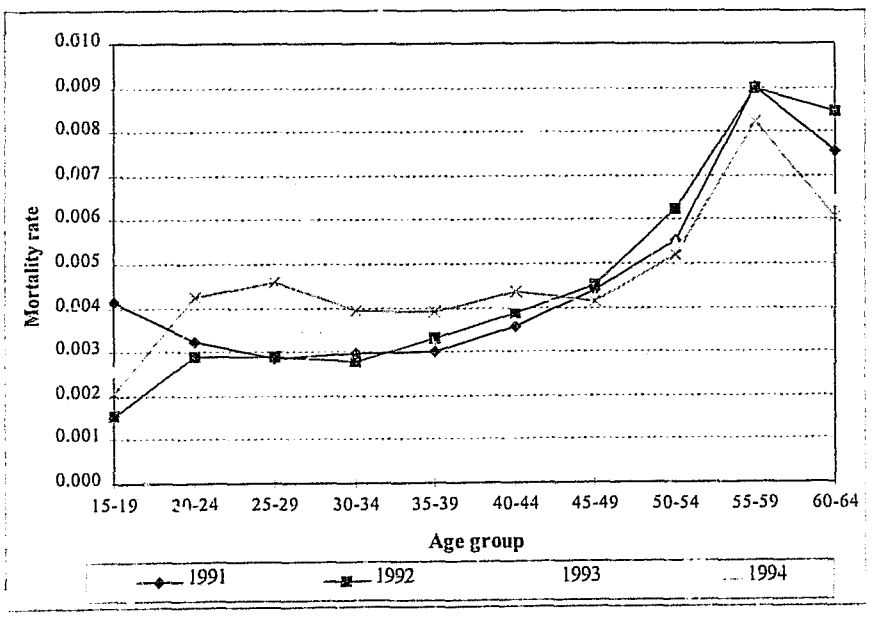


Figure 6.14 Estimated actual rates of mortality for: service occupations

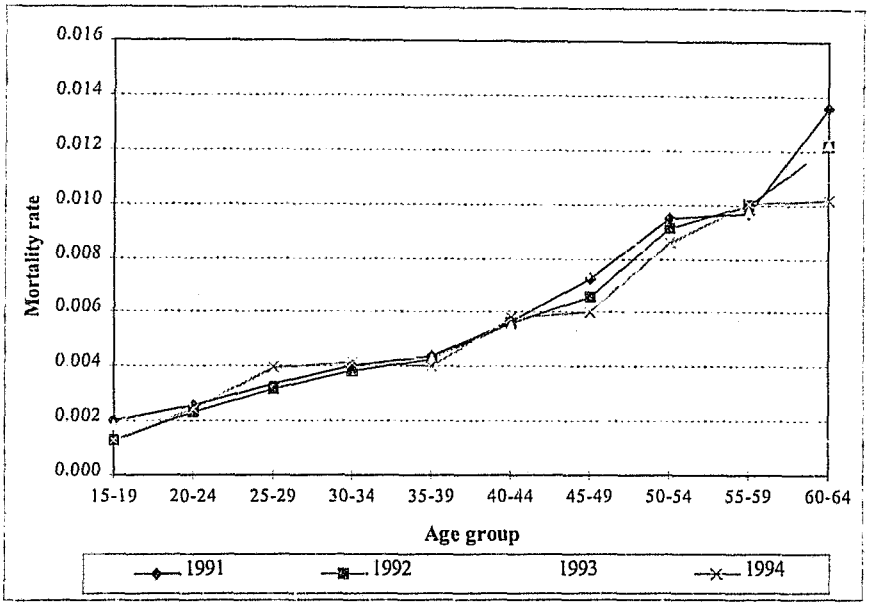


Figure 6.15 Estimated actual rates of mortality for: farming and related occupations

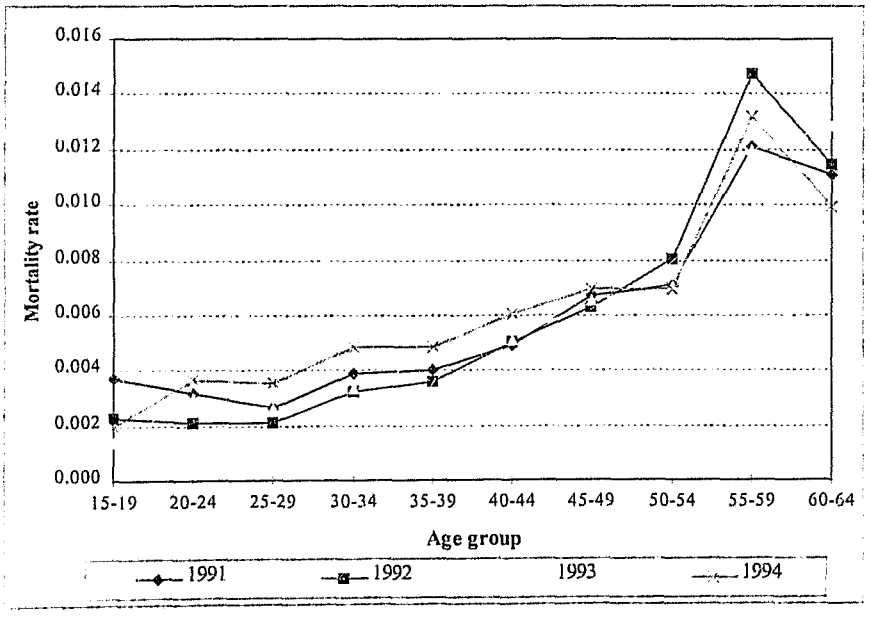


Figure 6.16 Estimated actual rates of mortality for: artisan, apprentice and related occupations

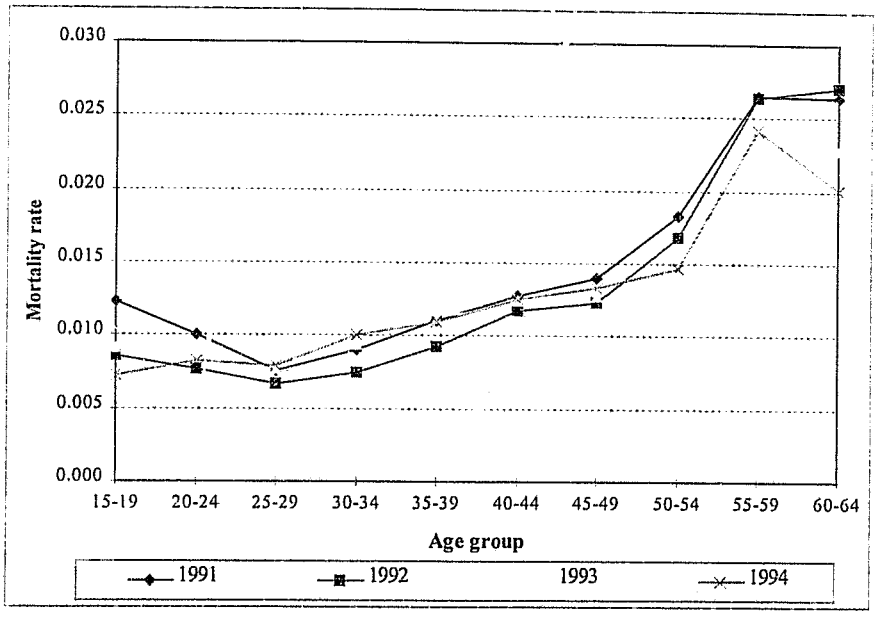


Figure 6.17 Estimated actual rates of mortality for: production supervisor, miner, quarry and related worker

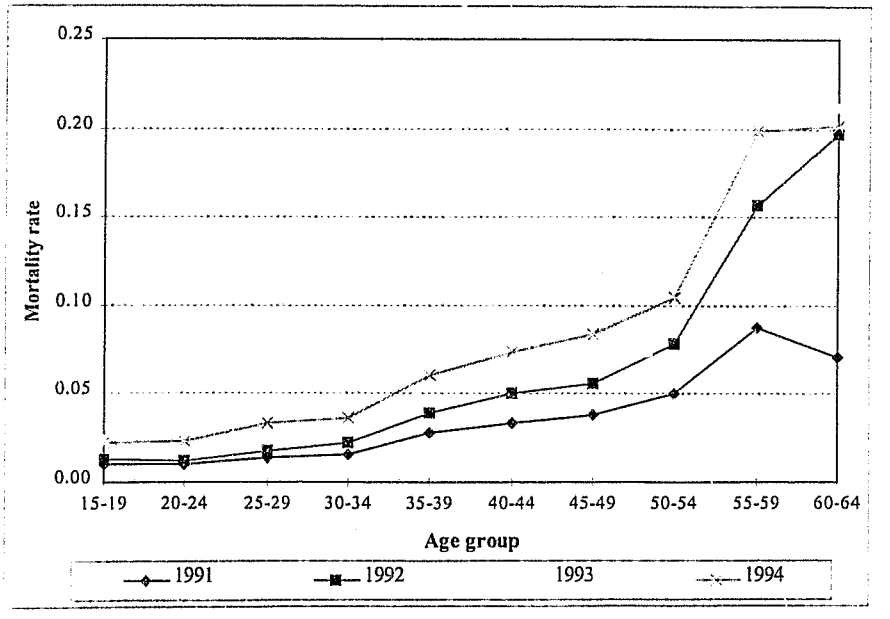


Figure 6.18 Estimated actual rates of mortality for: occupation unspecified and not elsewhere classified

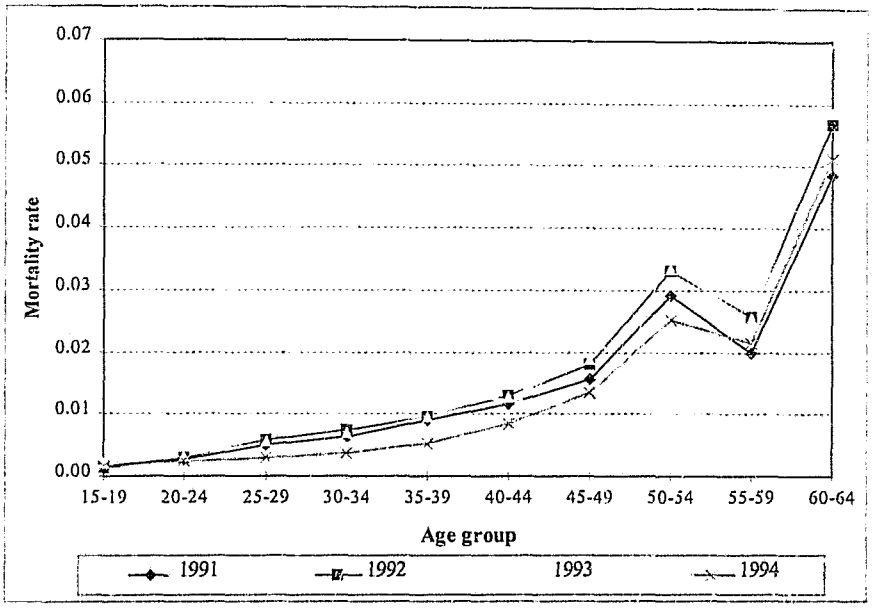


Figure 6.19 Estimated actual rates of mortality for: total not economically active

The following observations are made:

- The largest increases in mortality since 1991 occurred amongst the occupations which were unspecified and not elsewhere classified. This may be the result of a deterioration in the recording of occupation on the death certificate. However, it may be indicative of AIDS, especially if commercial sex workers have been classified under this occupation group.
- The second largest increases in mortality occurred within occupation groups 2 (managerial, executive and administrative occupations) and 4 (transport, delivery and communications occupations). Long-distance truck drivers are classified under occupation group 4, thus providing some justification for attributing the increase in mortality to AIDS. As discussed, there is evidence that managerial occupations exhibit higher HIV prevalence rates. The higher mortality rates can thus be justified, but may be due in part to the increase in the proportion of blacks within these traditionally "white" occupations.
- Other notable increases in mortality occurred within occupation groups 1 (professional, semi-professional and technical occupations), 3 (clerical and sales occupations), and 5 (service occupations) and 7 (artisan, apprentice and related occupations).
- Mortality has increased between the ages of 25 and 40 within occupation group 8 (production supervisors, miners, quarry and related workers), although to a relatively small degree. There is, however, evidence of high HIV prevalence rates within this group.

- Mortality appears to have decreased substantially between 1993 and 1994 within the economically inactive population (i.e. occupation group 10). This substantiates the claim that AIDS has the most significant impact upon the economically active population.

6.2.4 Results

The "excess" deaths arising in occupation group j from an increase in mortality between 1991 and year t (denoted ${}^{j}_{5}\overline{\overline{INCR}}_{t,x}$) were calculated for $t = 3, 4$ as shown in equation 6.10:

$$
\begin{aligned}
{}^{j}_{5}\overline{\overline{INCR}}_{t,x} &= ({}^{j}_{5}q_{t,x}{}^{A} - {}^{j}_{5}q_{1,x}{}^{A}).{}^{j}_{5}\overline{\overline{N}}_{t,x} && \text{for } x = 20, 25, 30, \ldots, 50 \\
{}^{j}_{5}\overline{\overline{INCR}}_{t,x} &= 0 && \text{elsewhere} \qquad \textit{(6.10)}
\end{aligned}
$$

No adjustment was made for annual changes in the level of the mortality curve for each occupation group j, due to a lack of information. The 1991 actual mortality rates were used as expected rates, even at the older ages when the rates became erratic. This was done because no standard tables exist by occupation group with which to smooth the actual rates.

Table 6.4 "Excess" adult deaths and percentage increase in deaths since 1991 by occupation group

Occupation group	1993		1994	
	Number	*Percentage*	Number	*Percentage*
1. Professional	532	*20%*	1 180	*42%*
2. Managerial	522	*21%*	1 033	*40%*
3. Clerical	185	*4%*	782	*16%*
4. Transport	539	*18%*	1 655	*55%*
5. Service	705	*11%*	1 643	*24%*
6. Farming	- 18	*0%*	- 167	*-3%*
7. Artisan	- 52	*-1%*	672	*17%*
8. Production	- 1 131	*-4%*	- 1 046	*-3%*
9. Unspecified	43 404	*99%*	55 168	*122%*
10. Not economically active	5 424	*0%*	- 9 547	*0%*
Total	50 111		51 371	

Since the size of the occupation groups differs, it is not possible to compare the "excess" deaths in absolute terms. The percentage increase in deaths above the expected level for occupation group j was calculated for t = 3, 4 as shown in equation 6.11:

$$\frac{100.{}_5D_{t,x}^{j=}}{{}_5q_{1,x}^{A}.({}_5N_{t,x}^{j=}+1/2.{}_5D_{t,x}^{j=})}-100 \qquad (6.11)$$

The "excess" deaths may be AIDS-related, and thus provide information on which occupation groups are most likely to be affected by the epidemic. The absolute deaths and the percentage increase in deaths appear in Table 6.4.

6.2.5 Discussion

As can be seen from the table, the total "excess" deaths are substantially larger than the number of AIDS-related deaths estimated in Chapter 5. This may be the result of the following:

- The level of the mortality curve may have increased within certain occupation groups. The "excess" deaths would thus incorporate changes in the level and shape of the mortality curve, in contrast to the estimates obtained in Chapter 5 which include only changes in the shape of the mortality curve.
- The proportion of blacks within certain occupation groups may have increased, possibly increasing the mortality and resulting in "excess" deaths.
- Other factors (such as violence) may have contributed to higher mortality within certain occupation groups, resulting in "excess" deaths.
- All four population groups were combined in the analysis. A worsening of asian, coloured and white mortality (which may incorporate some Pattern I AIDS-related deaths) would result in "excess" deaths which should not be compared against estimates of the number of black AIDS-related deaths.
- Deaths occurring amongst the economically active population are likely to be overstated, as mentioned in section 6.2.1.
- The proportion of the population within certain occupation groups may have increased over time. Equation 6.8 would thus result in underestimates of the population within those occupation groups, and mortality rates would be overstated.

As discussed in section 6.2.4, occupation groups 1, 2, 3, 4, 5, 7, and 8 are likely to contain some AIDS-related deaths. The total "excess" deaths from these groups are 1 300 and 5 919 in 1993 and 1994 respectively. This is substantially lower than the 12 018 and 22 608 AIDS-related deaths estimated for 1993 and 1994 respectively under Scenario 2. AIDS-related deaths would thus have been included within other occupation groups, with a high proportion perhaps in the unspecified category. Negative AIDS-related deaths are estimated for occupation

groups 7 and 8. This is probably due to unreliable data and random error.

The relative levels of AIDS-related deaths within different occupation groups can be determined by comparing the percentage increase in deaths since the level experienced in 1991. The largest increase in deaths occurred within occupation group 9. Large increases are also apparent within occupation groups 1 and 2, with the percentage increase doubling from 1993 to 1994. Occupation group 4 displayed a very large increase in deaths, especially in 1994 where it was second only to occupation group 9.

In conclusion, the analysis by occupation group yields unreliable results which do not reconcile with those obtained in section 5.7. However, there appears to be some indication of AIDS-related deaths occurring within occupation groups 1, 2, 3, 4, 5 and 9. The recording of occupation on the death certificate needs to be improved before reliable results can be produced from the analysis performed in section 6.2.

AIDS appears to be affecting skilled and semi-skilled workers. This has serious implications for the industries involved, as HIV-related absenteeism and morbidity will lower productivity and raise expenditure on recruitment, training, health care, death and funeral benefits (Foster, 1996).

6.3 Analysis of AIDS-Related Deaths in KwaZulu-Natal

The AIDS epidemic is more advanced in certain provinces as is evident from the provincial prevalence rates of antenatal clinic attendees (section 2.2.2). KwaZulu-Natal appears to be ahead of the rest of South Africa. This was confirmed by Doyle (1991). An analysis of mortality rates within this province would thus provide insight into the expected spread of the epidemic in the rest of the country.

The 1991, 1992, 1993 and 1994 death reports published by the CSS (1991 a, 1992, 1993 b and unpublished) provide data on the age and gender structure of reported deaths arising within KwaZulu-Natal. A breakdown by population group was not available, and it appeared possible to extract the reported black deaths from the aggregate reported deaths as in section 5.3.1. This required the assumption that the actual pattern of asian, coloured and white deaths arising in KwaZulu-Natal was equal to an expected pattern based on the adjusted SA 84-86 life tables. The validity of this assumption is questionable, and deaths from all population groups were therefore analysed (thus maintaining consistency with the analyses in sections 6.1 and 6.2).

The notation defined in section 5.2 will be used preceded by a K. For example,

$_{y-x}(KN)_{t,i,s,x}$

denotes the number of individuals in population group i living in KwaZulu-Natal, with gender s and aged between x and y on 30th June of year t.

6.3.1 Adjustment for Underreporting

The same estimates of the completeness of reporting ($_5\overline{C}_{t,s,x}$) were used as described in section 6.1.1. These estimates are likely to be incorrect for the population of KwaZulu-Natal, since the proportion of blacks is higher than in the country as a whole. The completeness of reporting is thus likely to be overstated.

The actual number of deaths occurring in year t amongst individuals living in KwaZulu-Natal, with gender s and aged between x and $x+5$ was estimated as follows:

$$_5\overline{(KD)}_{t,s,x} = \frac{_5\overline{(KD)}_{t,s,x}{}^{R}}{_5\overline{C}_{t,s,x}} \qquad (6.12)$$

Equation 6.12 is of the form of equations 6.2 and 6.5.

6.3.2 Estimation of the Population of KwaZulu-Natal

The 1991 census results provide an age and gender breakdown of the population living in Natal, as well as Statistical Regions 90, 91 and 92 (constituting the population of KwaZulu) (CSS, 1991 c). It was thus possible to deduce the size and demographic characteristics of the population of the existing KwaZulu-Natal province, as given by the 1991 census ($_5\overline{(KN)}_{1,s,x}$).

The population was classified into quinquennial age bands up to age 25, and into the following age groups thereafter: (25, 35), (35, 45), (45, 55), (55, 60), (60, 65), (65, ω). The estimates of the population aged between 25 and 55 were grouped into quinquennial age groups in line with the total population estimates from the 1991 census ($_{y-x}\overline{N}_{1,s,x}$), as follows:

$$_5\overline{(KN)}_{1,s,x} = {}_{b-a}\overline{(KN)}_{1,s,a} \cdot \frac{_5\overline{N}_{1,s,x}}{_{b-a}\overline{N}_{1,s,x}} \qquad (6.13)$$

where a and b are the lower and upper bounds of the age groups (25, 35), (35, 45) and (45, 55), and $x \geq a$ and $y \leq b$.

The proportion of the KwaZulu-Natal population with gender s which was aged between x and x+5 in 1991 ($_5(kp)_{s,x}$) was calculated as shown in equation 6.14:

$$_5(kp)_{s,x} = \frac{_5\overline{(KN)}_{1,s,x}}{\sum_{x=0}^{(\omega-5)/5} {}_5\overline{(KN)}_{1,s,5x}} \qquad (6.14)$$

These proportions were assumed to remain constant over time. This may be unrealistic, although it is fairly unlikely that a rapid change in age structure occurred between 1991 and 1994.

The CSS produced mid-year population estimates of the population of KwaZulu-Natal for 1992 to 1995 based on the 1991 census estimates (CSS: Personal communication). Let $\overline{EST}_{t,s}$ denote the estimate for gender s at the mid-point of year t.

The population of KwaZulu-Natal at the mid-point of year t (t = 1, 2, 3, 4), with gender s (s = 0, 1) and aged between x and x+5 was calculated as follows:

$$_5\overline{(KN)}_{t,s,x} = \overline{EST}_{t,s} \cdot {}_5(kp)_{s,x} \qquad (6.15)$$

6.3.3 Analysis of Mortality Rates in KwaZulu-Natal

The actual average probability for year t that an individual living in KwaZulu-Natal with gender s and aged between x and x+5 will die within one year, was calculated as shown in equation 6.16:

$$_5\overline{(kq)}_{t,s,x}{}^{A} = \frac{_5\overline{(KD)}_{t,s,x}}{_5\overline{(KN)}_{t,s,x} + 1/2 \cdot {}_5\overline{(KD)}_{t,s,x}} \qquad (6.16)$$

Equation 6.16 is analogous to equations 5.35 and 6.9.

The actual mortality curves are illustrated in Figures 6.20 and 6.21.

The level of mortality in KwaZulu-Natal appears to have been lowest in 1992 and increased thereafter. This pattern was found in sections 5.5 and 6.1.2, and is probably due to errors in the underlying data. AIDS-related deaths occurring in 1992 have therefore not be estimated. 1 427 politically motivated deaths occurred in KwaZulu-Natal during 1992, compared to 1 489 and 1 464 in 1993 and 1994 respectively (Jeffery, 1997). These statistics only partially explain the increase in mortality, and further investigation is required in order to fully justify the apparent increase.

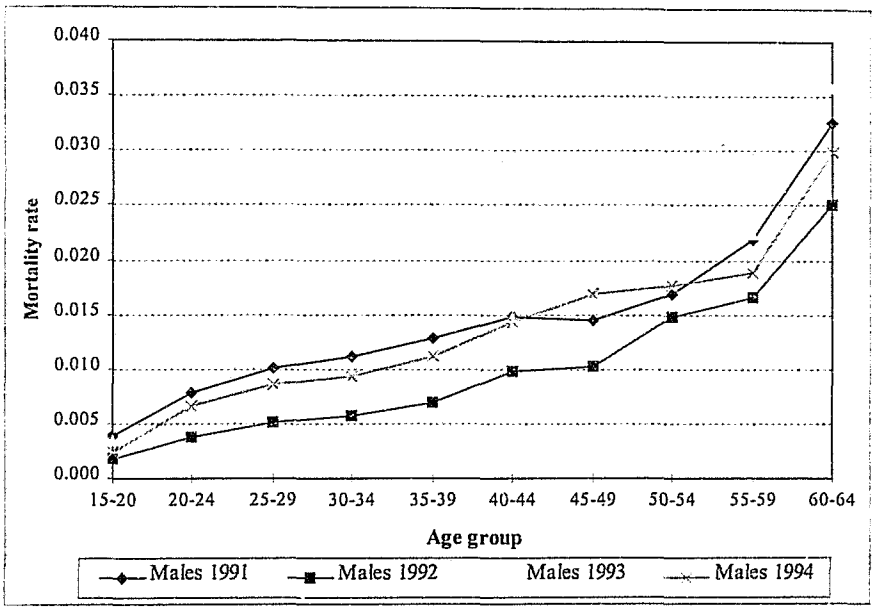


Figure 6.20 Estimated actual rates of mortality amongst males living in KwaZulu-Natal

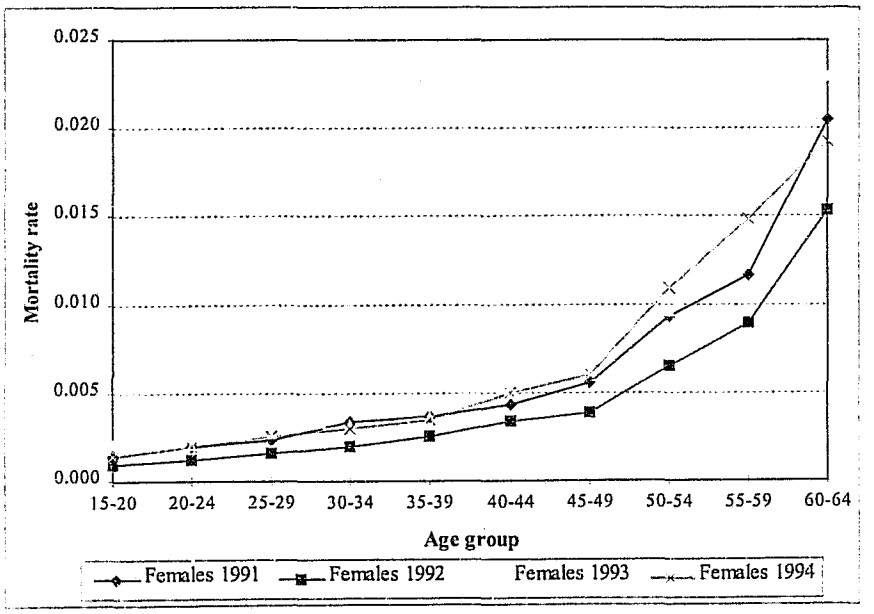


Figure 6.21 Estimated actual rates of mortality amongst females living in KwaZulu-Natal

6.3.4 Results

Changes in the *shape* rather than the *level* of the actual mortality curve may be attributed to AIDS. Some adjustment is therefore required before comparing 1993 and 1994 mortality rates with 1991 rates.

The 1991 actual mortality curves are viewed as "expected" mortality curves as in section 5.6. These expected mortality curves were adjusted upwards or downwards in 1993 and 1994 in order to negate any changes in the level.

The adjustment performed has been outlined in section 5.6.2. The KwaZulu-Natal estimates of the population (${}_5\overline{(KN)}_{t,s,x}$) and mortality rates (${}_5\overline{(kq)}_{t,s,x}{}^{A}$) were used, and the expected mortality rates were defined as the 1991 actual mortality rates for KwaZulu-Natal. These rates were adjusted for year t (t = 3, 4) and the iterative procedure outlined in section 5.7 was applied in order to remove distortions in the method of adjustment caused by changes in the shape of the actual mortality curve. No adjustment was made to the actual 1991 rates at the older ages, as random fluctuation appeared to be minimal relative to that of the total black population.

Results are shown in Table 6.5.

Table 6.5 Estimated number of AIDS-related deaths in KwaZulu-Natal

Age group	1993		1994	
	Males	Females	Males	Females
20-25	-483	3	-88	81
25-30	-503	55	-93	137
30-35	-529	-137	-122	-33
35-40	-399	-64	-36	31
40-45	-111	113	246	209
45-50	311	64	623	174
50-55	443	-72	345	344
Total	-1 271	-39	874	944

6.3.5 Discussion

The 1993 estimates are meaningless as they indicate a negative number of AIDS-

related deaths across almost all age bands. The 1994 estimates appear more reasonable.

When compared to Table 5.7, the above estimates suggest that 8% of the total male and female AIDS-related deaths occurred in KwaZulu-Natal during 1994. This may be compared to the proportion of HIV-positive women of childbearing age in 1994 who live in KwaZulu-Natal. By applying the prevalence rates from Tables 2.1 and 2.2 to estimates of the national and KwaZulu-Natal female populations aged between 15 and 50 (derived in sections 5.4 and 6.3.2 respectively), estimates of HIV-infected women can be obtained. Such calculation yields a proportion of 49%, which is substantially higher than the corresponding proportion of deaths. Further investigation is therefore necessary to justify the figure of 8%. The ratio of male to female deaths is 0,93:1 compared to a ratio of 1,12:1 estimated for the entire country.

The age distribution of the estimated deaths does not accord with the expected distribution illustrated in Figure 5.21. This was true for the estimates shown in Table 5.7. The distorting effect of these errors can be reduced by aggregating the data.

In conclusion, the estimated AIDS-related deaths which occurred in KwaZulu-Natal during 1994 do not reconcile with the national estimates. The KwaZulu-Natal estimates are likely to understate the epidemic in the province.

6.4 Summary

In contrast to the estimation method of Chapter 5, reported deaths from all population groups were analysed. These deaths were adjusted to allow for underreporting which varied by quinquennial age group, and were analysed in three ways:

1. Cause specific mortality rates (constituting a multiple decrement table) were calculated for 1991, 1992, 1993 and 1994. The increase in mortality rates within the age interval 20-55 was attributed to AIDS, and thus gave an indication of the commonly recorded causes of death amongst AIDS patients.

   Tuberculosis; other infectious and parasitic diseases; neoplasms; and endocrine, nutritional and metabolic diseases and immunity disorders were found to explain 15% and 13% of the incorrectly recorded male and female AIDS-related deaths occurring in 1994 amongst individuals aged 20-55. Together with the 11% correctly recorded AIDS-related deaths, this implies that 26% and 24% of male and female AIDS-related deaths occurring in 1994 can be explained.

2. Separate mortality rates were calculated for the ten occupation groups to

determine which occupations are likely to be significantly affected by AIDS. Few meaningful conclusions could be drawn. The largest annual increase in mortality between the ages of 20 and 55 was observed within occupation group 9 (occupation unspecified and not elsewhere classified). Increases were also observed within occupation groups 1 (professional, semi-professional and technical occupations), 2 (managerial, executive and administrative occupations), 3 (clerical and sales occupations), 4 (transport, delivery and communications occupations), and 5 (service occupations).

3. Mortality rates were calculated for KwaZulu-Natal which has historically led the AIDS epidemic in South Africa. Increases in mortality were attributed to AIDS. 8% of the 1994 AIDS-related deaths appear to have occurred in KwaZulu-Natal, although this figure is probably unreliable.

7

Calibration of the ASSA Model

The criterion involved in the calibration of a model is that projections correspond with new empirical data. The original ASSA model was calibrated to HIV prevalence data obtained from antenatal clinic surveys. In Chapter 5, empirical data on AIDS-related deaths were obtained. The ASSA model can thus be calibrated to these data but must maintain consistency with the HIV prevalence data. Such calibration is performed in this chapter, and is expected to validate or refine the model.

7.1 Adjustment of Demographic Assumptions Underlying the ASSA Model

The ASSA model requires a number of inputs regarding the demographic structure of the black population. In order to draw meaningful conclusions from the comparison of the model's projected AIDS-related deaths and the estimated actual AIDS-related deaths, the same population assumptions are required.

This section attempts to reconcile the demographic assumptions underlying the ASSA model to those assumed in the estimation of AIDS-related deaths (i.e. population estimates in section 5.4). In particular, the age and gender distribution of the population, the fertility assumptions and the size of the population in 1985 are adjusted. Normal mortality assumptions are not altered, and may differ from those implicit in the population estimates derived in section 5.4.

The resulting model will be referred to as the "adjusted ASSA model".

7.1.1 Adjustment of Age and Gender Distribution

The population projected by the ASSA model differs from that estimated in section 5.4 with respect to age and gender distribution. The assumed proportion of

the population at each age, and the male to female ratio at each age must thus be adjusted to reconcile with the age and gender distribution of the population estimated in section 5.4.

The age and gender distribution of the black population is likely to change during the period 1985 to 2005, unless the population is stable. For example, the apparent reduction in fertility rates reduces the proportion of children and leads to population ageing.

As discussed in section 4.3.2, the ASSA model requires values for p_x and σ_x for its age and gender structure. The model assumes that the underlying population is stable, and the values of p_x and σ_x were therefore adjusted so as to represent approximately average values for the period 1985 to 2005. Detailed 1991 census figures are readily available, and will thus be used to calculate adjusted values of p_x and σ_x. The 1991 structure may be roughly average if the rate of change in the population structure declines over the period 1985-2005.

In section 5.4.3, estimates of the black population in the former TBVC states were added to the 1991 census figures to derive estimates of the total black population of South Africa in 1991. These estimates were grouped into quinquennial age bands up to age 85, and a final open interval from age 85 upwards. The ASSA model, however, operates on individual ages. The estimates were thus distributed across the individual ages as follows, in the notation of section 5.2:

$$N_{1,4,s,x} = {}_5N_{1,4,s,a} \cdot \frac{{l_{4,s,x}}^S}{{}_5{l_{4,s,a}}^S} \qquad \text{where } a \le x \le a+4 \text{ and } x < 85$$

$$N_{1,4,s,x} = {}_{\omega-85}N_{1,4,s,85} \cdot \frac{{l_{4,s,x}}^S}{{}_{\omega-85}{l_{4,s,85}}^S} \qquad \text{where } x \ge 85 \qquad (7.1)$$

where ${l_{4,s,x}}^S$ refers to the life table derived by Dorrington, Martens and Slawski (1991) with the adjustments discussed in section 5.6.1.

p_x was thus derived as follows:

$$p_x = \frac{\sum_{s=0}^{1} N_{1,4,s,x}}{\sum_{s=0}^{1} {}_{\omega}N_{1,4,s,0}} \qquad \text{for } x < 90$$

$$p_x = \frac{\sum_{s=0}^{1} \omega - x N_{1,4,s,x}}{\sum_{s=0}^{1} \omega N_{1,4,s,0}} \qquad \text{for } x = 90 \qquad (7.2)$$

σ_x was derived as follows:

$$\sigma_x = \frac{N_{1,4,0,x}}{\sum_{s=0}^{1} N_{1,4,s,x}} \qquad \text{for } x < 90$$

$$\sigma_x = \frac{\omega - x N_{1,4,0,x}}{\sum_{s=0}^{1} \omega - x N_{1,4,s,x}} \qquad \text{for } x = 90 \qquad (7.3)$$

The new values of p_x and σ_x were used to adjust the ASSA model. The values are shown in Appendix D.

The model incorporates an assumption of the proportion of births which are male (γ_1). This was set to 0,52 in the original model. However, in order to be consistent with the estimated actual black population, this was changed to 0,515.

7.1.2 Adjustment of Total Fertility Rates

Recent evidence suggests that total fertility rates amongst black women have been declining (discussed in section 5.4.2). Fertility rates are assumed to be constant over the period 1985-2005 in the ASSA model, and appear to be overstated. Unless the fertility rates are reduced, the model will overstate the number of babies born each year, and will thus distort the age structure of the population.

The ASSA model requires an assumed total fertility rate for each risk group r (denoted TFR_r), as well as the proportion of the population falling in each risk group (denoted g_r). The values of g_r and TFR_r assumed under Scenario 500 of the ASSA model result in an overall total fertility rate of 4,11 (as explained in section 4.3.4).

Recent evidence indicates that the total fertility rate amongst black women was around 3,7 over the period 1990-1995, as discussed in section 5.4.2. If the decline in fertility rates slows, this value may represent a rough average for the period 1985-2005.

The assumed values of g_r were not altered since no further data were available. Adjusted total fertility rates applicable to each risk group r (denoted TFR_r') were calculated by solving equations 7.4 and 7.5 simultaneously. Equation 7.5 assumes that the ratios of the total fertility rates remain constant.

$$\sum_{r=2}^{5} g_r . TFR_r' = 3{,}7 \qquad (7.4)$$

$$\frac{TFR_r'}{TFR_2'} = \frac{TFR_r}{TFR_2} \qquad \text{for } r = 3, 4, 5 \qquad (7.5)$$

The adjusted total fertility rates were found to be 1,35, 2,48, 4,05 and 4,73 for the PRO, STD, RSK and NOT risk groups respectively. These fertility rates were used to adjust the ASSA model. The scaling factors which distribute the total fertility rates across individual ages (denoted t_x in section 4.3) were not altered.

7.1.3 Adjustment of Population Size

Once the age and gender distribution and the fertility assumptions had been adjusted, the size of the starting population in 1985 was varied to yield a population size in accordance with that derived in section 5.4.

A starting population (denoted *P(0)*) of 25 million in 1985 was assumed under Scenario 500 of the ASSA model. This figure was adjusted by a process of trial and error so that the size of the projected population aged between 0 and 75 in 1991 reconciled with the estimates derived in section 5.4.

The 1991 population projected by the model refers to the beginning of 1991, whereas the estimated population refers to the mid-point of 1991. The projected population in year t should thus be approximately half-way between the estimated population in year t-1 and t, assuming that the population grows uniformly throughout each year.

A starting population of 25,2 million was seen to produce reasonable results. This is shown in Table 7.1.

The unadjusted and adjusted starting populations are illustrated in Figures 7.1 and 7.2 respectively. By comparing the two figures, it is evident that the adjusted population assumes a higher proportion of young individuals and lower proportion of old individuals.

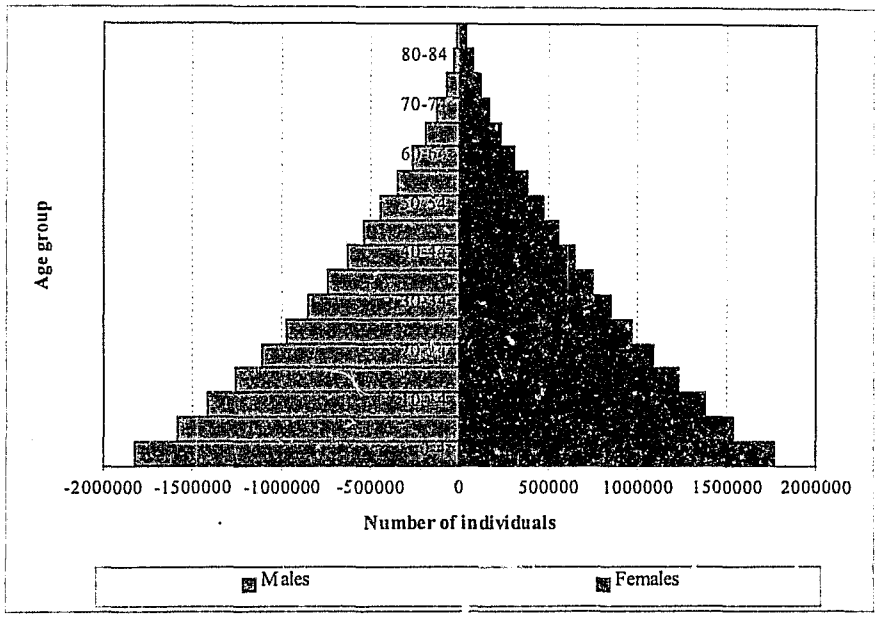


Figure 7.1 Population pyramid for the starting population in the unadjusted ASSA model

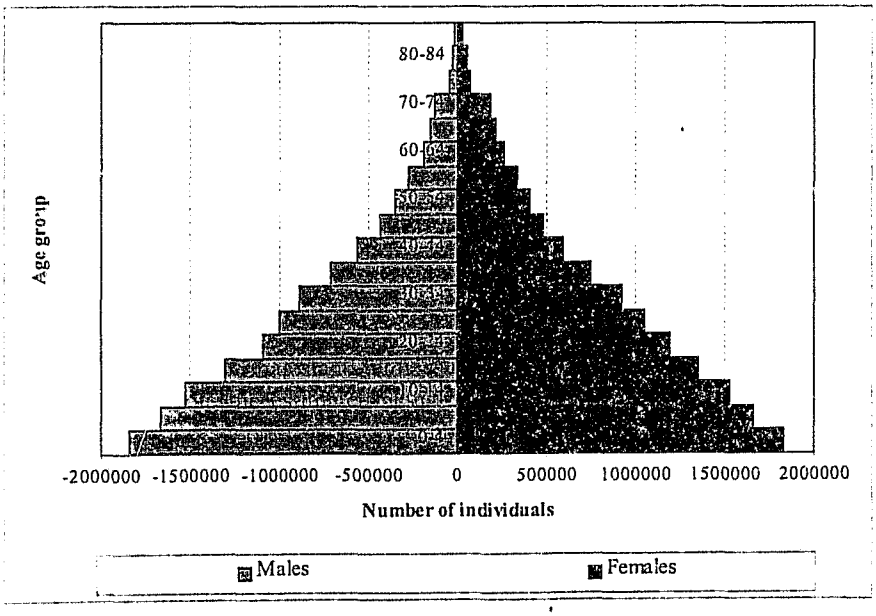


Figure 7.2 Population pyramid for the starting population in the adjusted ASSA model

Table 7.1 Comparison of the projected and estimated actual population aged 0 to 75 (in thousands), assuming a starting population of 25,2 million

Year	Males			Females		
	Projected	Scenario 2	Scenario 3	Projected	Scenario 2	Scenario 3
1991	13 714	13 690	13 690	14 290	14 399	14 399
1992	14 003	13 994	14 004	14 571	14 708	14 714
1993	14 296	14 336	14 350	14 855	15 040	15 050
1994	14 592	14 663	14 684	15 143	15 372	15 386

7.2 Calibration Against Estimated AIDS-Related Deaths

7.2.1 Comparison of the Adjusted ASSA Model and Empirical Data

The estimated number of AIDS-related deaths occurring amongst individuals aged 20-55 in 1993 and 1994 (derived in Chapter 5) can be compared with the number projected by the adjusted ASSA model. This comparison is shown in Table 7.2.

Table 7.2 Comparison of the estimated actual AIDS-related deaths occurring in 1993 and 1994 and those projected by the adjusted ASSA model

	Projected		Scenario 2		Scenario 3	
	1993	1994	1993	1994	1993	1994
Males	6 156	10 047	5 948	11 928	6 955	12 256
Females	5 558	9 583	6 070	10 680	6 185	10 369
Total	11 714	19 630	12 018	22 608	13 140	22 625

Assuming that the estimates of the number of AIDS-related deaths are correct, two observations can be made:

- The adjusted ASSA model understates the *total number* of AIDS-related deaths amongst individuals aged 20-55 by between 300 and 1 500 in 1993, and by about 3 000 in 1994.
- The adjusted ASSA model projects that 52,6% and 51,2% of AIDS-related deaths amongst individuals aged 20-55 were male in 1993 and 1994 respectively. As discussed in section 5.7.1, the male to female ratio of estimated actual AIDS-related deaths is is likely to be slightly greater than 1. The *gender distribution* of AIDS-related deaths projected by the ASSA model thus appears to be realistic.

The adjusted ASSA model will be calibrated so as to fit the estimated actual experience (shown in Table 7.2) in addition to the antenatal clinic data.

7.2.2 Calibration Method

The ASSA model's parameters can be changed in order to alter aspects of the projected epidemic. The modified parameters should obviously be justified by other empirical data.

The adjusted ASSA model was used as a starting point. The following inputs were identified as possibly influencing the projected number, and male to female ratio, of adult AIDS-related deaths (with notation in parentheses):

- The median term to death of an HIV-positive individual (m).
- The proportion of HIV-infected individuals surviving by duration d. This is determined by a reverse S-shaped curve (s_d) with two parameters: the median term to death (m), and k.
- The relative male to female infectivity factor (λ).
- The male and female indices of sexual activity. These indices are determined by smooth curves involving two parameters ($i_{s,x}$).
- The number of male and female HIV-infected immigrants in each year t (I_s).
- The average number of individuals in risk group r_j than an infected individual in risk group r_i would be expected to infect in one year, given that he/she only had contact with uninfected individuals (Ω_{r_i,r_j}).

The median term to death of an HIV-positive individual was assumed to be ten years under Scenario 500 of the ASSA model. However, evidence suggests that HIV-positive individuals do not live as long in Africa (Doyle, 1991). m may thus be set at a lower value.

It can be seen from equation 4.6 that $s_d = 0{,}5$ when $d = m$. The distribution of deaths about the median term to death is dependent on k as follows:

For $d < m$: A larger k results in a larger s_d.

For $d > m$: A larger k results in a smaller s_d.

Thus, a larger k implies that deaths are more concentrated around the median term to death, whereas a smaller k means that deaths are more evenly distributed throughout the 20 years.

k was arbitrarily defined as 0,5 in the original ASSA model, and may thus be changed for calibration purposes.

The relative male to female infectivity factor was set equal to two in the original ASSA model. Females were thus assumed to be twice as likely to be infected by an HIV-infected male per sexual contact than vice versa. This parameter is subject to much uncertainty, as values as far apart as two and ten have been documented (Schoub, 1994; *AIDS Analysis Africa (Southern Africa Edition)*, 1995 and Solomon, 1995).

The male and female curves of sexual activity at age x involve two parameters - a_s and b_s - as shown in equation 4.8.

Values of a_s and b_s were arbitrarily defined, although the curves were designed to reflect the fact that female sexual activity tends to peak at an earlier age than for males. There is thus no reason why a_s and b_s cannot be changed.

The ASSA model assumes that HIV-infected immigrants started the HIV epidemic in 1985. However, the chosen number of such immigrants was arbitrary, and there is thus no reason why the values of I_s could not be altered.

The values of Ω_{r_i,r_j} assumed in the original ASSA model were based on empirical evidence, as discussed in section 4.3.7. The parameters are therefore believed to be accurate, and adjustment cannot be justified at this stage.

Calibration of the ASSA model involves the adjustment of a large number of parameters to nine data points (the seven antenatal clinic results and the estimated male and female AIDS-related deaths in 1994). There is thus no unique set of parameters which would best calibrate the model, and hence no optimisation method can be used.

The parameters m, k, λ, a_s, b_s and I_s were thus adjusted simultaneously using a method of trial and error in order to reconcile the projected and estimated AIDS-related deaths shown in Table 7.3. Moreover, correspondence with the antenatal clinic data was maintained.

A satisfactory model was chosen by inspection.

7.2.3 The Calibrated Model

The satisfactory model was defined as follows:

- $m = 9$.
- k = 0,54. The adjusted and unadjusted curves of s_d are shown in Figure 7.3. Since k has been increased, deaths are more concentrated around the median .
- λ was left unchanged at 2.
- a_0 was left unchanged, and b_0 = 0,003. Similarly, a_1 was left unchanged, and b_1 = 0,006.

  The adjusted male and female indices are illustrated in Figure 7.4. The adjustment implies that the peak level of sexual activity amongst males is lower and occurs at a later age than before, but declines less rapidly after reaching the peak. The adjustment results in a higher peak level of sexual activity for females, with this peak occurring at an earlier age than before. Sexual activity declines more rapidly thereafter.
- I_s was left unchanged.

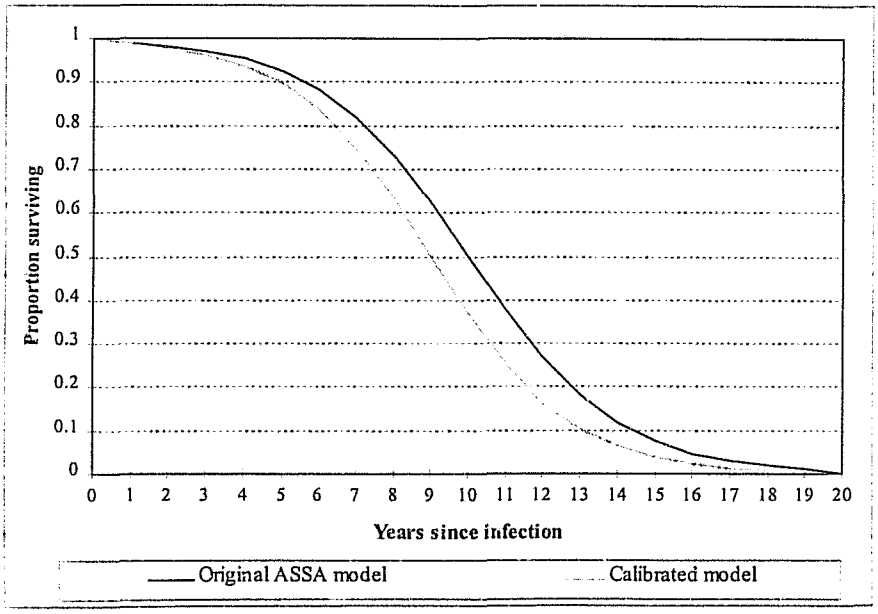


Figure 7.3 Curves of the proportion of HIV-infected individuals surviving by duration in the original and calibrated ASSA models

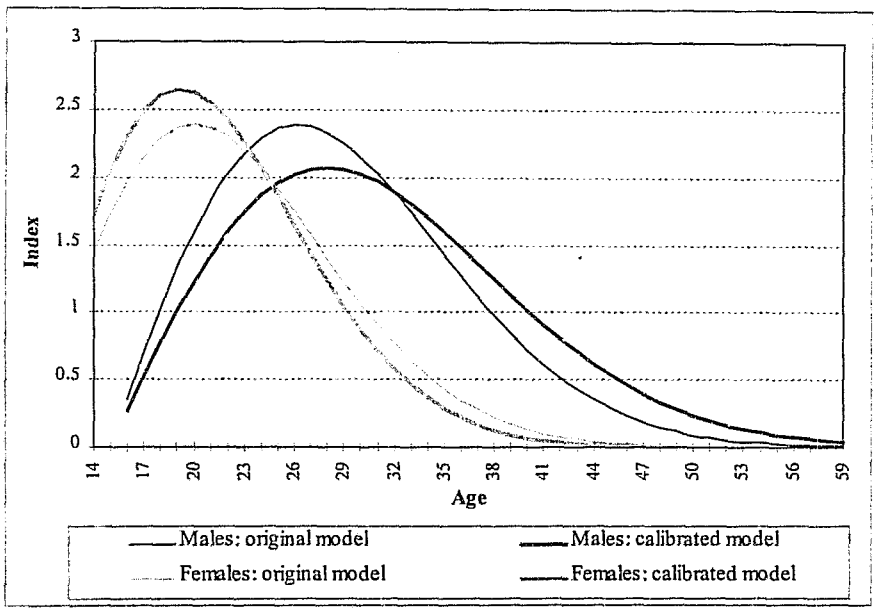


Figure 7.4 Sexual activity indices assumed in the original and calibrated ASSA models

The results of the calibration are shown in Table 7.3. The projected values are slightly higher than the estimated actual values. 54,2% and 52,6% of the projected AIDS-related deaths are male in 1993 and 1994 respectively. The calibration thus appears to have rectified the inaccuracies identified in section 7.2.1.

Figure 7.5 compares the antenatal clinic data and the HIV prevalence of pregnant women as projected by the calibrated model. Figure 7.5 can be compared with Figure 4.2, which examines the fit of the original model. It can be seen that the calibrated model still represents the results fairly well.

Table 7.3 Comparison of the estimated actual AIDS-related deaths occurring in 1993 and 1994 and those projected by the calibrated ASSA model

	Projected		Scenario 2		Scenario 3	
	1993	1994	1993	1994	1993	1994
Males	7 163	12 056	5 948	11 928	6 955	12 256
Females	6 062	10 860	6 070	10 680	6 185	10 369
Total	13 225	22 916	12 018	22 608	13 140	22 625

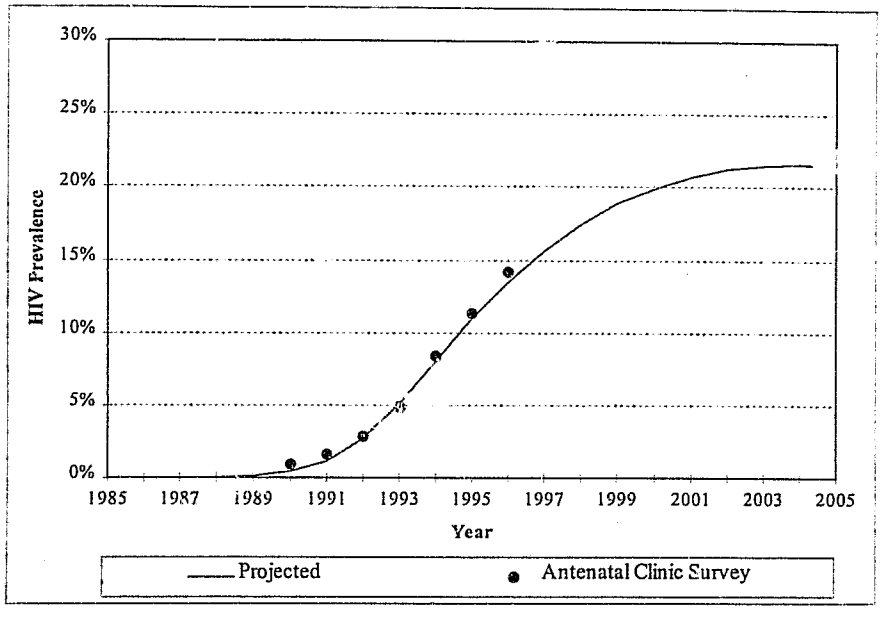


Figure 7.5 Comparison of the antenatal clinic data published by the Department of Health, and the HIV prevalence of pregnant women as projected by the calibrated model

7.3 Goodness-of-Fit Tests

A chi-squared goodness-of-fit test (Clarke and Cooke, 1992:289-314) was performed to test how well the calibrated ASSA model fits the estimated actual deaths (from all causes, including AIDS) within each age group in 1994 (Test 1). For the purpose of comparison, a similar test was performed to test the fit of the adjusted ASSA model in 1994 (Test 2).

- Let $s = 0, 1$ denote males and females respectively, as before.
- Let ${}_5O_{s,x}$ denote the estimated actual deaths occurring during 1994 amongst individuals with gender s and aged between x and $x+5$.
- Let ${}_5E_{s,x}^{C}$ denote the deaths occurring during 1994 amongst individuals with gender s and aged between x and $x+5$, as projected by the *calibrated* ASSA model.
- Let ${}_5E_{s,x}^{A}$ denote the deaths occurring during 1994 amongst individuals with gender s and aged between x and $x+5$, as projected by the *adjusted* ASSA model.

7.3.1 Test 1

The following null hypothesis was tested:
H_0: The estimated actual deaths in each age group are equal to the deaths projected by the *calibrated* ASSA model
against the following alternate hypothesis:
H_1: The estimated actual deaths in each age group are not equal to the deaths projected by the *calibrated* ASSA model.

A separate test statistic was calculated for males and females, as follows:

$$X_s^2 = \sum_{x=20}^{26} \frac{({}_5O_{s,5(x-20)+20} - {}_5E_{s,5(x-20)+20}{}^{C})^2}{{}_5E_{s,5(x-20)+20}{}^{C}} \qquad (7.6)$$

and $X_s^2 \sim \chi^2_5$ under H_0.

The calculated values of the test statistic are:

$X^2_0 = 4556$
$X^2_1 = 363119$

which are highly significant.

H_0 is thus rejected, implying that the calibrated model fits the estimated actual deaths very poorly.

7.3.2 Test 2

The following null hypothesis was tested:
H_0: The estimated actual deaths in each age group are equal to the deaths projected by the *adjusted* ASSA model

against the following alternate hypothesis:
H_1: The estimated actual deaths in each age group are not equal to the deaths projected by the *adjusted* ASSA model.

A separate test statistic was calculated for males and females, as follows:

$$X_s^2 = \sum_{x=20}^{26} \frac{({}_5O_{s,5(x-20)+20} - {}_5E_{s,5(x-20)+20}{}^{A})^2}{{}_5E_{s,5(x-20)+20}{}^{A}} \qquad (7.7)$$

and $X_s^2 \sim \chi^2_5$ under H_0.

The calculated values of the test statistic are:

$X^2_0 = 15005$
$X^2_1 = 138620$

which are highly significant.

H_0 is thus rejected, implying that the adjusted model also fits the estimated actual deaths very poorly.

7.3.3 Conclusion

The adjusted and calibrated ASSA models fit the estimated actual deaths very poorly. This is very unsatisfactory. The problem lies with the age distribution of the estimated actual deaths rather than that predicted by the models. The age distribution of the estimated actual AIDS-related deaths was noted to be unrealistic (section 5.7.1), and poor fits were thus anticipated.

By comparing Tables 7.2 and 7.3, it can be seen that the calibrated model represents the total estimated actual AIDS-related deaths better than the adjusted model.

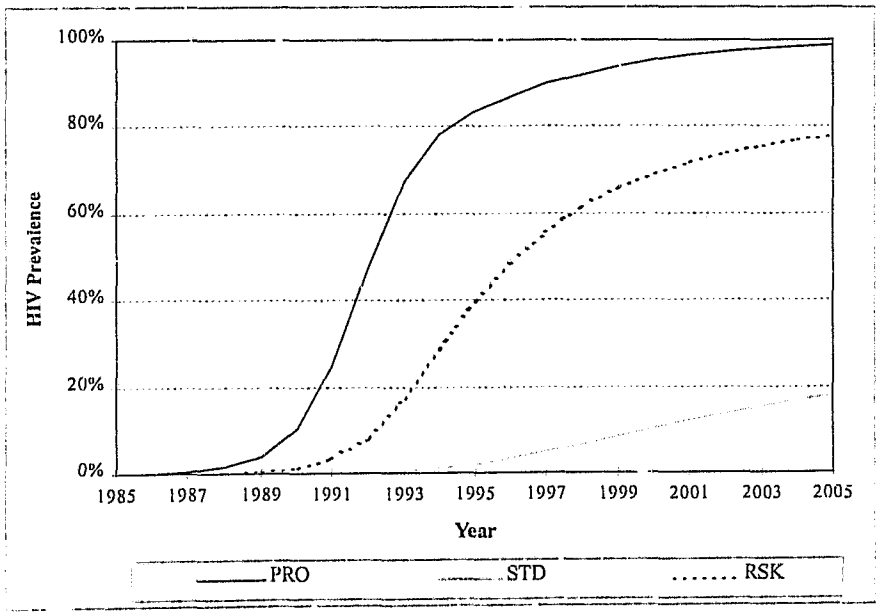


Figure 7.6 Revised projected HIV prevalence rates amongst females aged 15-45 in the PRO, STD and RSK risk groups

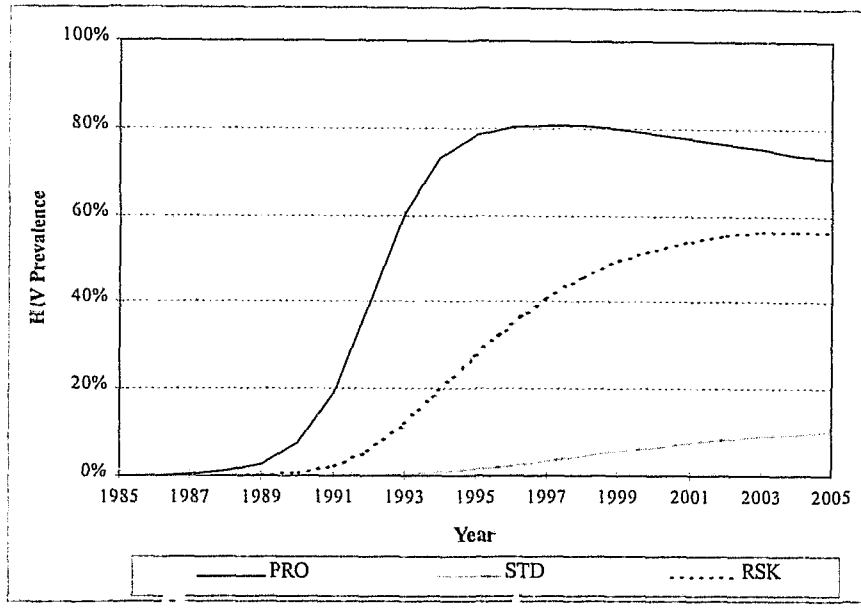


Figure 7.7 Revised projected HIV prevalence rates amongst males aged 15-45 in the PRO, STD and RSK risk groups

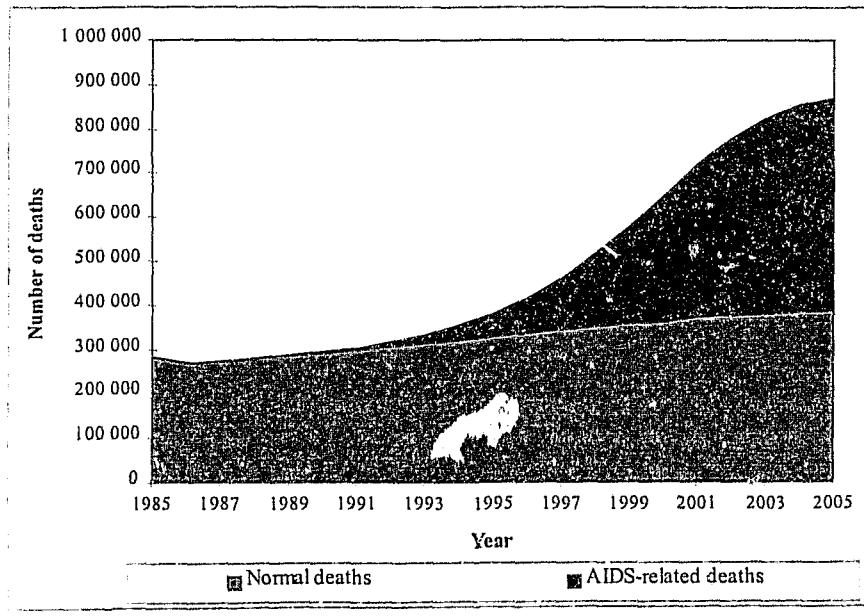


Figure 7.8 Revised projected normal and AIDS-related deaths

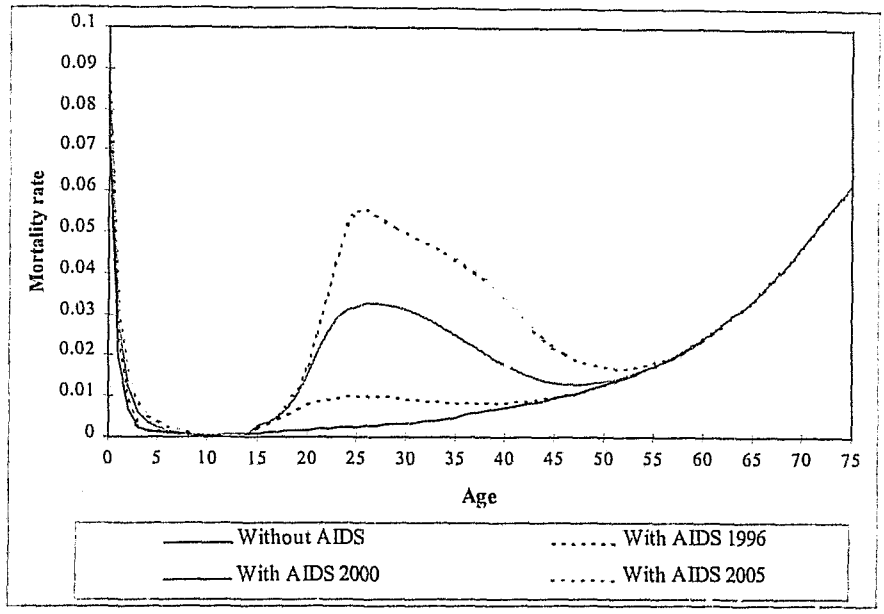


Figure 7.9 Revised projected female mortality curves, with and without the effect of AIDS

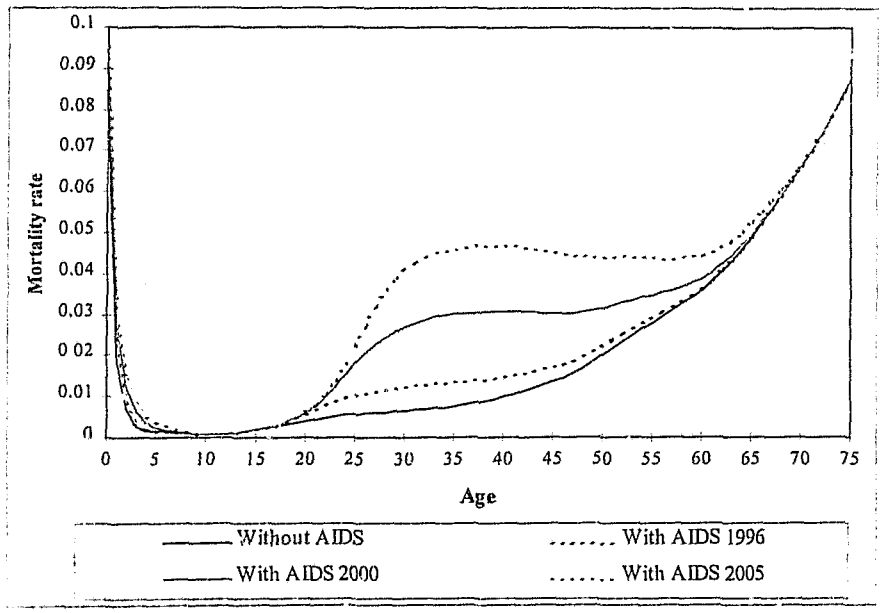


Figure 7.10 Revised projected male mortality curves, with and without the effect of AIDS

7.4 Revised Projections

The calibrated model can be used to provide revised projections of the future course of the AIDS epidemic in South Africa. Table 7.4 and Figures 7.6 to 7.10 present these revised projections for selected years, and projected age and sex specific mortality rates for the years 2000 and 2005 are shown in Appendix E.

Table 7.4 Revised projected statistics indicating the course of the AIDS epidemic in South Africa

	1991	1994	1996	2000	2005
HIV-positive males	101 695	676 365	1 135 060	1 780 400	1 904 207
HIV-positive females	137 439	877 779	1 450 912	2 155 105	2 178 978
Total HIV-positive	239 135	1 554 145	2 585 972	3 935 505	4 083 185
Deaths					
Normal deaths	299 798	321 671	335 809	363 123	386 230
AIDS-related deaths	3 802	31 446	78 588	286 348	482 731
Cumulative AIDS-related deaths	5 913	64 329	193 889	983 180	3 137 296
AIDS-related deaths					
Males aged 20-55	1 561	12 056	29 186	109 279	187 639
Females aged 20-55	1 286	10 860	29 980	132 748	230 069
Total aged 20-55	2 847	22 916	59 166	242 028	417 708
Crude extra mortality rate for total black population	0,0001	0,0010	0,0025	0,0086	0,0139

7.5 Discussion

7.5.1 Comparison with the Original ASSA Model

The original and calibrated ASSA models both appear to fit the antenatal clinic survey data fairly adequately, as can be seen from Figures 4.2 and 7.5. The original model fits slightly better according to the least squares criterion, but the

differences do not appear significant.

As can be seen by comparing Tables 4.2 and 7.4, the number of HIV-positive males projected by the calibrated model is larger than that projected by the original model until 2005. The number of infected females is larger until 2000. The total number of HIV-infected individuals projected by the calibrated model increases from approximately 91 000 in 1990 to 4,1 million in 2005. The original model projects that 88 000 individuals were infected in 1990, increasing to 4,4 million in 2005. The male to female ratios of the number of HIV-infected individuals projected by the calibrated model are slightly larger than those projected by the original model.

The projected prevalence rates amongst various subgroups of the population can be compared by comparing Figures 4.6 and 4.7 with Figures 7.6 and 7.7 respectively. The calibrated model tends to project lower peak prevalence rates, with the peaks occurring earlier. For example, the HIV prevalence of the female population aged 15-45 is estimated to have levelled-off by the year 2005, reaching a peak prevalence of around 27% in 2003. The original model projects a prevalence rate of around 29% in 2005, and the prevalence is estimated to be increasing. The HIV prevalence rate amongst males aged 15-45 is projected to be 20% and 22% in 2005 by the calibrated and original models respectively.

The calibrated model projects a larger number of AIDS-related deaths than the original model. Accumulated AIDS-related deaths projected by the calibrated model are expected to reach the 3 million mark in 2005, compared to a figure of 2,6 million projected by the original model.

The mortality rates projected by the calibrated model are thus affected by AIDS to a greater extent than those projected by the original model (as can be seen from Figures 4.4, 4.5, 7.9 and 7.10). Moreover, the crude rate of extra mortality due to AIDS is estimated to be 0,1% higher in 2005.

In summary, the calibrated ASSA model predicts higher AIDS-related mortality than the original model. HIV prevalence rates are predicted to peak at a lower level, in general.

7.5.2 Comparison with Empirical Data

The ratio of male to female HIV-positive individuals projected by the calibrated model increases over time. The ratio is 0,74:1 in 1991, 0,78:1 in 1996 and 0,87:1 in 2005. Similarly, the projected ratios of male to female prevalence rates amongst the population aged 15-45 increases from 0,75:1 in 1991 to 0,76:1 in 2005. Various ratios have been presented in the literature.

A commonly quoted ratio of male to female HIV-positive individuals is 0,73:1.

This ratio was estimated by the Blood Transfusion Services of South Africa (Department of Health, 1996), and is dependent on assumptions of the population size and gender structure.

The HIV prevalence amongst new blood donors throughout South Africa would be a more meaningful figure. This was determined by the Blood Transfusion Services of South Africa in half-year intervals from July 1987 to December 1989 (Schoub *et al*, 1990). The male to female ratios of HIV prevalence amongst black donors varied between 0,60:1 to 1:1.

The ratios projected by the calibrated model appear to reconcile with both sources of data mentioned above. Further research is necessary before any definite conclusions can be drawn.

The number of HIV-positive individuals projected by the calibrated model accords with 1996 estimates derived by the Department of Health (1996/1997). In particular, the projected number of infected males falls between the best and high estimates of 1 018 429 and 1 167 588 respectively. The projected number of infected females falls between the best estimate of 1 395 108 and the high estimate of 1 599 435. The total number of projected HIV-positive individuals is very close to the Department of Health's best estimate. This justifies the model's results.

7.6 Summary

The ASSA model was calibrated against the AIDS-related deaths which are estimated to have occurred in 1993 and 1994. Adequate representation of the antenatal clinic HIV prevalence data was maintained.

The demographic assumptions of the model were adjusted to reconcile with those used to estimate AIDS-related deaths. In particular, the age and gender distribution and the size of the population was altered. Fertility rates were reduced to accord with recent evidence.

Calibration was performed using the method of "trial-and-error" since parameters were more plentiful than data points. Four parameters were adjusted:

- The median duration to death of an HIV-positive individual was reduced from 10 to 9 years.
- A parameter of the survival probabilities by duration was changed, resulting in a greater concentration of deaths around the median.
- Parameters of the male and female sexual activity indices were changed so as to decrease the peak level of sexual activity for males, and increase it for females.

The calibrated model projects a greater number of AIDS-related than the original model. Projected prevalence rates are lower, and peak earlier.

Certain parameters will need to be changed over time as new empirical data become available, and as sexual behaviour alters. Calibration should thus be an ongoing process.

8

Application to the Insurance Industry

The ASSA model is a valuable tool in the determination of a set of mortality assumptions appropriate to the individual life office. This is illustrated in this chapter, and the importance of model accuracy is emphasised.

Life assurers test all new applicants for HIV, and can thus ensure that HIV-positive lives are rejected at the outset. However, the life office may still be exposed to AIDS-related mortality in two ways: firstly, policyholders who were underwritten before the introduction of HIV testing may have been HIV-positive, and secondly, policyholders who produce a negative test result at the outset may become infected at a later stage.

8.1 Preliminaries

Since the ASSA model projects the spread of the AIDS epidemic amongst the black population, it can only be used to determine mortality assumptions for black policyholders. The rest of this chapter thus applies to black assured lives.

The notation of Chapter 4 is used where appropriate. It is assumed that HIV-testing was introduced at the start of year T, and that policies were effected at the start of year τ.

Assured lives are assumed to be in risk group r (where $r = 2, 3, 4, 5, 6$), and aged x last birthday (where $x = 14, 15, 16, ...$). Policyholders are assumed to effect policies between the ages of 14 and 60 (i.e. whilst in risk group $r = 2, 3, 4, 5$). It is assumed that a life office is able to determine the proportion of its assured lives in risk group r at policy inception (denoted $\phi_r(\tau)$),

where $\sum_{r=2}^{5} \phi_r(\tau) = 1$.

The mortality rate which should be assumed by a life office in year t in respect of policies written at time τ, for assured lives with gender s and aged x last birthday at time t, will be denoted ${}_{\tau}q_{s,x}(t)'$, where $t > \tau$. The effect of underwriting for diseases other than AIDS will not be considered, and the rates thus represent ultimate rates.

8.2 Policies Written Prior to HIV-Testing

For policies written in year τ, where $\tau < T$, the life office is unsure about the proportion of policyholders which were HIV-positive at inception. The proportion of *policyholders* in risk group r which were HIV-positive at time τ is assumed to be equal to the proportion of the *population* in risk group r which were HIV-positive at time τ. Policyholders are thus assumed to experience population mortality.

The ultimate mortality rates which should be assumed by a life office during year t in respect of policies effected in year τ (for $\tau < T$) are given by equation 8.1:

$$
{}_{\tau}q_{s,x}(t)' = \sum_{r=2}^{5} \phi_r(\tau).\left\{ \frac{C_{s,x,r}(t).q_{s,x} + \sum_{i=max(0;t-20)}^{t-1} F_{s,x,r,t-i-1}(t).{q_{t-i-1}}^{A}}{C_{s,x,r}(t) + \sum_{i=max(0;t-20)}^{t-1} F_{s,x,r,t-i-1}(t)} \right\} \quad \text{for } x < 60
$$

$$
{}_{\tau}q_{s,x}(t)' = \frac{C_{s,x,6}(t).q_{s,x} + \sum_{i=max(0;t-20)}^{t-1} \left\{ \left(\sum_{r=2}^{5} F_{s,x,r,t-i-1}(t) \right).{q_{t-i-1}}^{A} \right\}}{C_{s,x,6}(t) + \sum_{i=max(0;t-20)}^{t-1} \sum_{r=2}^{5} F_{s,x,r,t-i-1}(t)} \quad \text{for } x \geq 60
$$

(8.1)

8.3 Policies Written Subsequent to HIV-Testing

It is assumed that only uninfected individuals will be accepted for insurance. (HIV-positive individuals may, in fact, be accepted for specific policies. These lives can be assumed to be subject to the AIDS mortality rates $q_d{}^A$, where d is the duration since infection.)

The ultimate rates which should be assumed by a life office during year t in respect of policies effected in year τ (where $\tau \geq T$) are calculated as follows:

$$ {}_{\tau}q_{s,x}(t)' = \sum_{r=2}^{5} \phi_r(\tau).\left\{ \frac{C_{s,x,r}(t).q_{s,x} + \sum_{i=\tau}^{min(t-1;19)} F_{s,x,r,t-i-1}(t).q_{t-i-1}^{A}}{C_{s,x,r}(t) + \sum_{i=\tau}^{min(t-1;19)} F_{s,x,r,t-i-1}} \right\} \quad \text{for } x < 60 $$

$$ {}_{\tau}q_{s,x}(t)' = \frac{C_{s,x,6}(t).q_{s,x} + \sum_{i=\tau}^{min(t-1;19)} \left\{ \left(\sum_{r=2}^{5} F_{s,x,r,t-i-1}(t) \right).q_{t-i-1}^{A} \right\}}{C_{s,x,6}(t) + \sum_{i=\tau}^{min(t-1;19)t-1} \sum_{r=2}^{5} F_{s,x,r,t-i-1}(t)} \quad \text{for } x \geq 60 $$

(8.2)

8.4 Example

8.4.1 Statement of the Problem

Consider the following hypothetical situation:

A particular life office, XYZ Life, started testing for HIV on 1 January 1990. The actuaries assume that 33% of its black policyholders are in the NOT risk group at policy inception, 50% in the RSK risk group, and 17% in the STD risk group. (These proportions are in line with the proportions used by the Actuarial Society of South Africa to derive recommended bases of extra mortality for group lives, under Professional Guidance Note 105 (ASSA, 1995)).

The actuaries of XYZ Life need to estimate mortality rates to be used in 1998 in respect of two cohorts of policies:

a) Whole life assurances effected on 1 January 1987 by males aged 45 last birthday.
b) Annuities effected on 1 January 1995 by females aged 30 last birthday.

8.4.2 Solution

$\phi_r(\tau) = \phi_r$ for all τ, and $\phi_2 = 0$, $\phi_3 = 0{,}17$, $\phi_4 = 0{,}5$ and $\phi_5 = 0{,}33$.

a) Males aged 45 last birthday at 1 January 1987 are aged 56 last birthday at 1 January 1998. The mortality rate ${}_2q_{0,56}(13)'$ is thus required.

Equation 8.1 yields ${}_2q_{0,56}(13)' = 0{,}029495$ ♦

b) Females aged 30 last birthday on 1 January 1995 are aged 33 last birthday on 1 January 1998. The mortality rate ${}_{10}q_{1,33}(13)'$ is thus required.

Equation 8.2 yields ${}_{10}q_{1,33}(13)' = 0{,}004463$ ♦

The solution to part (a) yielded ${}_{2}q_{0,56}(13)' = 0{,}029495$. The corresponding normal (non-AIDS) population mortality rate is slightly lower at 0,029446. XYZ Life should increase the normal mortality rate by 0.2% to allow for the impact of AIDS. The mortality rate for the population aged 56 last birthday at 1 January 1988 is predicted to equal the normal (non-AIDS) mortality rate. The assured lives actually experience heavier mortality than the population, which is unusual. This phenomenon is likely to be the result of values of ϕ_r which are different to g_r (where g_r is defined in section 4.3.1).

The solution to part (b) yielded ${}_{10}q_{1,33}(13)' = 0{,}004463$, compared to the corresponding normal mortality rate of 0,004199. XYZ Life should increase the normal mortality rate by 6% in order to allow for the effect of AIDS on this cohort of policies. The mortality rate for the females in the general population aged 33 last birthday at 1 January 1998 is predicted to equal 0,016210. The population of assured lives are thus expected to experience lighter mortality than the general population, as expected.

8.5 Discussion

The above example highlights the importance of correctly allowing for the effect of AIDS on mortality rates. The increase over the normal mortality rates is dependent on the age and gender of the cohort of lives, the mix of risk groups within the particular cohort, and the policy inception date.

The results are sensitive to the model's projections, and the need for model accuracy is therefore stressed. The model should be calibrated as soon as new empirical evidence becomes available in order to refine the projections.

8.6 Summary

Life assurers are exposed to AIDS-related mortality in two ways: (i) individuals who effected policies prior to the introduction of HIV-testing may have been HIV-positive at the outset; and (ii) individuals who were tested (and were thus HIV-negative) when the policy was effected may have become infected at a later stage.

The ASSA model can be used to estimate mortality rates for a group of assured

lives with a particular risk profile. Separate techniques are required for policies effected prior and subsequent to the introduction of HIV-testing.

9

Conclusion

Models of the AIDS epidemic are generally based on readily available HIV prevalence data. Projected HIV prevalence levels should therefore accord with empirical evidence. In contrast, the projection of AIDS-related deaths is subject to considerable uncertainty as little empirical data exist.

The AIDS-related deaths estimated in this dissertation are an attempt to address this problem. The estimation process was complicated by considerable underreporting of deaths, late reporting of deaths, inconsistency of reporting practices over time, and controversial adjustments to population enumeration. Approximations and adjustment methods were applied, resulting in unreliable estimates. As these data inadequacies are (hopefully) resolved with time, and as the number of AIDS-related deaths increases, less approximation should be required.

The 1993 and 1994 estimates are believed to be relatively dependable. It is suggested that between 10 000 and 15 000 and between 20 000 and 25 000 AIDS-related deaths occurred amongst the black population aged 20 to 55 during 1993 and 1994 respectively. Males constitute more than 50% of the deaths.

The age structure of the estimates of AIDS-related deaths did not correspond with previous empirical evidence, but is believed to be a result of poor data quality and random error. Aggregating across the age bands is likely to reduce the random error. Further research will need to be undertaken to determine the age structure of AIDS-related deaths.

A large number of AIDS-related deaths are likely to be incorrectly recorded with respect to cause of death. An analysis of cause-specific mortality rates was conducted. Large mortality increases were attributed to tuberculosis - a common ailment amongst AIDS patients. Significant mortality increases were also identified amongst six particular occupation groups.

The ASSA model estimates that 17 243 AIDS-related deaths occurred amongst

the black population aged 20 to 55 during 1994. 53% were projected to be amongst males. The projections appear to be lower than the empirical estimates derived in this dissertation.

The ASSA model was calibrated to project the total male and female AIDS-related deaths estimated in this dissertation, whilst maintaining consistency with the antenatal HIV prevalence data. Chi-squared goodness-of-fit tests indicate that the calibrated model fits the empirical data poorly within individual age bands. This was anticipated due to the unrealistic age structure of the empirical estimates.

The calibrated model projects that 1 million AIDS-related deaths will have occurred by the year 2005, in contrast to the 2,6 million projected by the original ASSA model. Projected mortality rates are affected. The epidemic is predicted to be less severe amongst the male population.

The calibrated model may not be a true representation of reality. It does, however, accord with what empirical evidence is available. Calibration should be a continuous process in order to incorporate behavioural changes and new data. This will benefit all users of the model, and particularly the life insurance industry which bases future mortality assumptions on realistic predictions of the spread of the epidemic.

APPENDICES

Table A1 Estimated number of reported black deaths under Scenario 2

Age group	1991 Total	1991 Male	1991 Female	1992 Total	1992 Male	1992 Female
0-5	17 819	9 504	8 315	16 855	9 221	7 635
5-10	1 271	775	496	945	584	361
10-15	1 308	789	520	1 017	598	419
15-20	2 821	2 022	799	2 025	1 361	664
20-25	5 059	3 792	1 266	3 806	2 722	1 084
25-30	5 851	4 450	1 401	4 806	3 480	1 327
30-35	6 312	4 665	1 647	5 472	3 841	1 631
35-40	6 407	4 609	1 798	5 723	4 047	1 676
40-45	6 167	4 504	1 663	6 041	4 214	1 827
45-50	5 736	4 197	1 539	5 492	3 851	1 641
50-55	6 883	4 528	2 355	6 755	4 543	2 212
55-60	5 790	3 440	2 350	6 364	3 957	2 407
60-65	7 665	4 421	3 245	7 982	4 643	3 339
65-70	6 383	3 806	2 577	7 264	4 292	2 971
70-75	7 623	4 060	3 563	9 347	5 111	4 236
75-80	4 338	2 392	1 946	5 114	2 781	2 333
80-85	3 039	1 417	1 622	4 161	1 930	2 231
85+	5 433	2 038	3 395	6 314	2 386	3 928
Total	105 906	65 409	40 496	105 483	63 561	41 922

Age group	1993 Total	1993 Male	1993 Female	1994 Total	1994 Male	1994 Female
0-5	17 637	9 477	8 160	16 279	8 745	7 534
5-10	1 486	863	623	1 488	844	644
10-15	1 570	948	622	1 486	882	604
15-20	3 797	2 489	1 307	3 597	2 377	1 220
20-25	6 362	4 709	1 653	7 098	5 084	2 014
25-30	7 260	5 291	1 969	8 402	5 947	2 456
30-35	7 727	5 503	2 224	8 743	6 123	2 620
35-40	7 804	5 588	2 216	3 512	5 821	2 691
40-45	7 677	5 352	2 325	8 814	5 965	2 849
45-50	6 960	4 868	2 091	8 267	5 752	2 515
50-55	8 000	5 275	2 725	8 549	5 564	2 984
55-60	7 380	4 531	2 849	8 275	5 283	2 992
60-65	8 977	4 969	4 008	9 111	4 913	4 198
65-70	8 402	4 967	3 435	9 518	5 260	4 258
70-75	9 046	5 042	4 004	8 783	4 912	3 871
75-80	6 548	3 350	3 199	7 771	3 814	3 958
80-85	4 200	1 903	2 297	5 554	2 582	2 972
85+	6 324	2 292	4 032	7 175	2 510	4 665
Total	127 158	77 417	49 740	137 422	82 377	55 045

Table B1 Estimated number of actual black deaths under Scenario 2

Age group	1991 Total	1991 Male	1991 Female	1992 Total	1992 Male	1992 Female
20-25	11 354	8 214	3 140	7 334	4 968	2 366
25-30	13 113	9 638	3 475	9 246	6 352	2 894
30-35	14 187	10 104	4 083	10 570	7 011	3 559
35-40	14 441	9 982	4 459	11 044	7 387	3 656
40-45	13 880	9 755	4 125	11 678	7 692	3 986
45-50	12 906	9 090	3 816	10 611	7 030	3 581
50-55	15 646	9 807	5 839	13 119	8 292	4 827
55-60	13 279	7 451	5 827	12 474	7 224	5 251
Total	108 805	74 039	34 765	86 076	55 957	30 120

Age group	1993 Total	1993 Male	1993 Female	1994 Total	1994 Male	1994 Female
20-25	12 264	8 795	3 469	13 198	9 261	3 937
25-30	14 014	9 884	4 131	15 633	10 833	4 801
30-35	14 945	10 278	4 667	16 275	11 155	5 121
35-40	15 087	10 438	4 649	15 865	10 605	5 260
40-45	14 876	9 997	4 879	16 435	10 867	5 569
45-50	13 482	9 094	4 388	15 394	10 478	4 917
50-55	15 571	9 853	5 717	15 971	10 137	5 834
55-60	14 442	8 462	5 979	15 472	9 624	5 848
Total	114 682	76 802	37 880	124 244	82 958	41 286

Table C1 Estimated black population aged 0 to 75 under Scenario 2

Age group	1991 Total	1991 Male	1991 Female	1992 Total	1992 Male	1992 Female
0-5	4 147 831	2 079 493	2 068 337	4 177 993	2 106 791	2 071 202
5-10	3 738 228	1 874 757	1 863 471	3 803 282	1 907 170	1 896 112
10-15	3 443 404	1 723 040	1 720 364	3 499 381	1 751 723	1 747 659
15-20	2 998 541	1 474 188	1 524 353	3 083 221	1 521 145	1 562 076
20-25	2 577 603	1 226 823	1 350 780	2 652 502	1 269 563	1 382 939
25-30	2 319 516	1 127 447	1 192 069	2 358 778	1 138 156	1 220 622
30-35	2 049 536	1 004 029	1 045 506	2 089 921	1 018 756	1 071 165
35-40	1 652 759	805 302	847 457	1 718 131	835 046	883 085
40-45	1 303 579	633 970	669 609	1 359 739	658 628	701 111
45-50	1 034 199	490 566	543 633	1 075 385	510 273	565 113
50-55	860 205	400 108	460 098	880 894	409 158	471 736
55-60	680 398	309 282	371 116	702 879	319 554	383 325
60-65	504 074	214 937	289 137	522 475	224 270	298 205
65-70	419 240	179 802	239 438	421 935	178 811	243 124
70-75	359 669	146 522	213 147	355 566	144 751	210 815
Total	28 088 782	13 690 266	14 398 516	28 702 083	13 993 795	14 708 288

Age group	1993 Total	1993 Male	1993 Female	1994 Total	1994 Male	1994 Female
0-5	4 232 005	2 143 864	2 088 142	4 297 088	2 184 658	2 112 430
5-10	3 862 699	1 939 755	1 923 944	3 921 638	1 972 573	1 949 065
10-15	3 558 177	1 781 767	1 776 410	3 616 227	1 811 668	1 804 560
15-20	3 164 363	1 566 135	1 598 228	3 238 106	1 606 307	1 631 798
20-25	2 733 102	1 316 158	1 416 944	2 809 221	1 358 973	1 450 248
25-30	2 409 008	1 158 486	1 250 522	2 460 466	1 180 429	1 280 037
30-35	2 133 774	1 035 827	1 097 947	2 174 454	1 050 227	1 124 227
35-40	1 782 061	864 546	917 515	1 837 963	888 521	949 442
40-45	1 420 211	686 486	733 725	1 478 060	712 248	765 812
45-50	1 121 850	533 048	588 802	1 168 192	551 745	613 447
50-55	908 100	421 921	486 179	936 573	435 [illegible]	501 578
55-60	726 210	330 097	396 113	748 277	[illegible]	408 487
60-65	543 396	234 759	308 636	562 913	244 412	318 501
65-70	428 144	180 238	247 906	434 963	181 996	252 968
70-75	351 956	142 731	209 225	350 329	140 977	209 352
Total	29 376 056	14 335 818	15 040 238	30 034 471	14 662 519	15 371 952

Table D1 Age and gender distributions in the unadjusted ASSA model

Age	Age distribution	Gender distribution	Age	Age distribution	Gender distribution
0	0.0325	0.509	46	0.0090	0.491
1	0.0292	0.509	47	0.0087	0.490
2	0.0281	0.509	48	0.0084	0.490
3	0.0273	0.508	49	0.0081	0.489
4	0.0267	0.508	50	0.0078	0.488
5	0.0261	0.508	51	0.0076	0.486
6	0.0255	0.508	52	0.0073	0.485
7	0.0249	0.508	53	0.0070	0.484
8	0.0244	0.507	54	0.0067	0.483
9	0.0238	0.507	55	0.0064	0.481
10	0.0233	0.507	56	0.0062	0.480
11	0.0228	0.507	57	0.0059	0.478
12	0.0223	0.507	58	0.0056	0.477
13	0.0218	0.506	59	0.0054	0.475
14	0.0213	0.506	60	0.0051	0.473
15	0.0209	0.506	61	0.0048	0.470
16	0.0204	0.506	62	0.0046	0.468
17	0.0199	0.505	63	0.0044	0.465
18	0.0195	0.505	64	0.0041	0.463
19	0.0190	0.505	65	0.0039	0.459
20	0.0186	0.504	66	0.0036	0.456
21	0.0181	0.504	67	0.0034	0.452
22	0.0177	0.504	68	0.0032	0.448
23	0.0172	0.503	69	0.0030	0.443
24	0.0168	0.503	70	0.0028	0.443
25	0.0164	0.503	71	0.0026	0.431
26	0.0160	0.502	72	0.0024	0.423
27	0.0156	0.502	73	0.0022	0.414
28	0.0152	0.502	74	0.0020	0.404
29	0.0148	0.501	75	0.0018	0.390
30	0.0144	0.501	76	0.0017	0.380
31	0.0140	0.500	77	0.0015	0.370
32	0.0136	0.500	78	0.0014	0.360
33	0.0133	0.500	79	0.0012	0.350
34	0.0129	0.499	80	0.0011	0.340
35	0.0126	0.499	81	0.0009	0.330
36	0.0122	0.498	82	0.0008	0.320
37	0.0119	0.497	83	0.0007	0.310
38	0.0116	0.497	84	0.0006	0.300
39	0.0112	0.496	85	0.0005	0.290
40	0.0109	0.496	86	0.0005	0.280
41	0.0106	0.495	87	0.0004	0.270
42	0.0103	0.494	88	0.0003	0.260
43	0.0099	0.494	89	0.0003	0.250
44	0.0096	0.493	90	0.0002	0.240
45	0.0093	0.492			

Table D2 Age and gender distributions in the adjusted ASSA model

Age	Age distribution	Gender distribution	Age	Age distribution	Gender distribution
0	0.0327	0.504	46	0.0074	0.476
1	0.0291	0.501	47	0.0073	0.474
2	0.0283	0.501	48	0.0072	0.473
3	0.0280	0.501	49	0.0070	0.471
4	0.0279	0.501	50	0.0063	0.470
5	0.0264	0.502	51	0.0062	0.468
6	0.0264	0.502	52	0.0061	0.465
7	0.0263	0.502	53	0.0059	0.463
8	0.0263	0.501	54	0.0058	0.460
9	0.0262	0.501	55	0.0051	0.460
10	0.0243	0.501	56	0.0049	0.457
11	0.0243	0.500	57	0.0048	0.454
12	0.0243	0.500	58	0.0046	0.452
13	0.0242	0.500	59	0.0045	0.448
14	0.0242	0.500	60	0.0038	0.432
15	0.0212	0.492	61	0.0037	0.429
16	0.0212	0.492	62	0.0036	0.426
17	0.0211	0.492	63	0.0034	0.423
18	0.0211	0.491	64	0.0033	0.420
19	0.0210	0.491	65	0.0033	0.436
20	0.0183	0.478	66	0.0031	0.433
21	0.0182	0.477	67	0.0030	0.429
22	0.0182	0.476	68	0.0028	0.424
23	0.0181	0.475	69	0.0026	0.420
24	0.0180	0.474	70	0.0029	0.416
25	0.0165	0.488	71	0.0027	0.411
26	0.0164	0.487	72	0.0025	0.407
27	0.0163	0.486	73	0.0024	0.402
28	0.0162	0.485	74	0.0022	0.397
29	0.0161	0.484	75	0.0012	0.413
30	0.0146	0.492	76	0.0011	0.407
31	0.0145	0.491	77	0.0010	0.401
32	0.0144	0.490	78	0.0009	0.393
33	0.0143	0.489	79	0.0008	0.385
34	0.0142	0.488	80	0.0009	0.391
35	0.0119	0.489	81	0.0008	0.382
36	0.0118	0.488	82	0.0007	0.371
37	0.0116	0.487	83	0.0006	0.360
38	0.0115	0.486	84	0.0005	0.348
39	0.0114	0.486	85	0.0004	0.401
40	0.0094	0.488	86	0.0004	0.386
41	0.0093	0.487	87	0.0003	0.371
42	0.0092	0.486	88	0.0003	0.355
43	0.0091	0.485	89	0.0002	0.339
44	0.0089	0.484	90	0.0002	0.321
45	0.0075	0.477			

Table E1 Mortality rates projected by the calibrated ASSA model - 2000

Age	Males	Females	Age	Males	Females
0	0.093708	0.086962	38	0.030541	0.020520
1	0.027566	0.028317	39	0.030616	0.019102
2	0.012500	0.012593	40	0.030657	0.017761
3	0.006390	0.005951	41	0.030660	0.016528
4	0.003616	0.003509	42	0.030617	0.015432
5	0.002611	0.002395	43	0.030521	0.014498
6	0.001975	0.001516	44	0.030380	0.013748
7	0.001441	0.001049	45	0.030246	0.013202
8	0.001156	0.000872	46	0.030178	0.012873
9	0.000843	0.000600	47	0.030235	0.012763
10	0.000789	0.00058	48	0.030472	0.012871
11	0.000836	0.000604	49	0.030924	0.013184
12	0.000971	0.000659	50	0.031533	0.013667
13	0.001187	0.000742	51	0.032230	0.014283
14	0.001477	0.000850	52	0.032945	0.014994
15	0.001832	0.002808	53	0.033614	0.015768
16	0.002243	0.003540	54	0.034195	0.016586
17	0.002958	0.004994	55	0.034737	0.017469
18	0.003833	0.007257	56	0.035310	0.018455
19	0.004887	0.010603	57	0.035981	0.019578
20	0.006188	0.015176	58	0.036812	0.020878
21	0.007837	0.020575	59	0.037860	0.022381
22	0.009940	0.025541	60	0.038755	0.024055
23	0.012512	0.028974	61	0.040367	0.025858
24	0.015346	0.030944	62	0.042272	0.027745
25	0.018080	0.031998	63	0.044475	0.029676
26	0.020486	0.032504	64	0.046977	0.031630
27	0.022523	0.032623	65	0.049735	0.033649
28	0.024195	0.032430	66	0.052693	0.035797
29	0.025629	0.031937	67	0.055809	0.038134
30	0.026834	0.031191	68	0.059064	0.040717
31	0.027832	0.030221	69	0.062458	0.043579
32	0.028639	0.029068	70	0.066014	0.046639
33	0.029269	0.027776	71	0.069767	0.049795
34	0.029735	0.026392	72	0.073768	0.052944
35	0.030065	0.024944	73	0.078073	0.055985
36	0.030291	0.023464	74	0.082742	0.058908
37	0.030441	0.021981	75	0.087802	0.062047

Table E2 Mortality rates projected by the calibrated ASSA model - 2005

Age	Males	Females	Age	Males	Females
0	0.094587	0.088853	38	0.046707	0.038228
1	0.029725	0.030474	39	0.046732	0.035976
2	0.014640	0.014733	40	0.046692	0.033558
3	0.008356	0.007918	41	0.046579	0.031065
4	0.005330	0.005223	42	0.046382	0.028602
5	0.004051	0.003836	43	0.046089	0.026271
6	0.003119	0.002660	44	0.045702	0.024148
7	0.002276	0.001885	45	0.045262	0.022281
8	0.001707	0.001423	46	0.044827	0.020689
9	0.001169	0.000926	47	0.044451	0.019375
10	0.000960	0.000751	48	0.044191	0.018349
11	0.000912	0.000680	49	0.044086	0.017613
12	0.000999	0.000687	50	0.044088	0.017159
13	0.001196	0.000751	51	0.044138	0.016966
14	0.001477	0.000850	52	0.044176	0.017005
15	0.001832	0.002894	53	0.044150	0.017242
16	0.002243	0.003663	54	0.044026	0.017643
17	0.003000	0.005222	55	0.043862	0.018213
18	0.003921	0.007746	56	0.043729	0.018969
19	0.005022	0.011677	57	0.043699	0.019926
20	0.006389	0.017473	58	0.043835	0.021108
21	0.008158	0.025334	59	0.044198	0.022526
22	0.010507	0.034732	60	0.044225	0.024140
23	0.013615	0.044154	61	0.045196	0.025902
24	0.017572	0.051534	62	0.046486	0.027765
25	0.022244	0.055402	63	0.048110	0.029683
26	0.027207	0.055833	64	0.050076	0.031632
27	0.031888	0.054421	65	0.052343	0.033649
28	0.035769	0.052781	66	0.054856	0.035797
29	0.038773	0.051368	67	0.057560	0.038134
30	0.041026	0.050108	68	0.060409	0.040717
31	0.042732	0.048894	69	0.063393	0.043579
32	0.044026	0.047664	70	0.066582	0.046639
33	0.044988	0.046399	71	0.070058	0.049795
34	0.045675	0.045072	72	0.073886	0.052944
35	0.046143	0.043641	73	0.078110	0.055985
36	0.046443	0.042050	74	0.082751	0.058908
37	0.046620	0.040254	75	0.087804	0.062047

REFERENCES

1. *AIDS Analysis Africa (Southern Africa Edition)* (1995). Vol. 5, No. 6.

2. *AIDS Analysis Africa (Southern Africa Edition)* (1996). Vol. 7, No. 2, pp. 12.

3. *AIDS Scan* (1997). Vol. 9, pp. 9.

4. Anderson, R. M., Blythe, S. P., Gupta, S. and Konings, E. (1989). The transmission dynamics of the human immunodeficiency virus type 1 in the male homosexual community in the United Kingdom: the influence of changes in sexual behaviour. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, Vol. 325, pp. 45-98.

5. ASSA (1995). *Professional Guidance Note 105: Recommended AIDS Extra Mortality Bases.* Actuarial Society of South Africa, Cape Town.

6. ASSA (1996). *The ASSA AIDS Model.* Unpublished, Website http://www.und.ac.za/und/eco/eru/eru.htm.

7. Bah, S. and Kleinschmidt, I. (1997). History of birth and death statistics in South Africa and a summary of the latest reports on births and deaths. *Vital Registration Meeting*, Department of Health. Pretoria, 11 June 1997.

8. Barnett, T. and Blaikie, P. (1992). *AIDS in Africa: Its Present and Future Impact.* Belhaven Press, London.

9. Bennett, N. G. and Horiuchi, S. (1981). Estimating the completeness of death registration in a closed population. *Population Index*, Vol. 47(2), pp. 207-221.

10. Bongaarts, J. (1989). A model of the spread of HIV infection and the demographic impact of AIDS. *Statistics in Medicine*, Vol. 8, pp. 103-120.

11. Booth, H. (1979). *The Estimation of Fertility from Incomplete Cohort Data by*

Means of the Transformed Gompertz Model. PhD Thesis, London School of Hygiene and Tropical Medicine, University of London.

12. Bos, E. and Bulatao, R. A. (1992). The demographic impact of AIDS in sub-Saharan Africa. Short- and long-term projections. *International Journal of Forecasting*, Vol. 8, pp. 367-384.

13. Botha, J. L. and Bradshaw, D. (1985). African vital statistics - a black hole? *South African Medical Journal*, Vol. 67, pp. 977-981.

14. Bowie, D., Cairns, A., Macdonald, A. and Speed, C. (1996). Models, useful models, and reality. *The Actuary*, Dec 1996, pp. 27-28.

15. Bradshaw, D., Dorrington, R. E. and Sitas, F. (1992). The level of mortality in South Africa in 1985 - what does it tell us about health? *South African Medical Journal*, Vol. 82, pp. 237-240.

16. Brass, W. (1974). Perspectives in population prediction: illustrated by the statistics of England and Wales. *Journal of the Royal Statistical Society A*, Vol. 137, Part 4.

17. Brass, W. (1981). The use of the relational Gompertz model to estimate fertility. *Proceedings of the International Union for the Scientific Study of Population Conference*, Manila, Liege.

18. Broomberg, J., Steinberg, M., Masobe, P. and Behr, G. (1991). *The Economic Impact of the AIDS Epidemic in South Africa.* Paper No. 23, Part II. The Centre for Health Policy, Department of Community Health, Medical School, University of the Witwatersrand, Johannesburg.

19. Brown, P. (1990). Africa's growing AIDS crisis. *New Scientist*, Vol. 128, No. 1743, pp. 38-41.

20. *Business Day* (1997). 6% have AIDS virus, survey finds. 25 April 1997.

21. Calitz, J. M. (1991). *Southern African Population Projections: 1995-2005.* Centre for Information Analysis, Development Bank of Southern Africa.

22. Calitz, J. M. (1996). *Population of South Africa: Updated Estimates, Scenarios and Projections 1990-2020.* Centre for Policy and Information Development Paper No. 109, Development Bank of Southern Africa.

23. Central Statistical Office of Zimbabwe (1985). *1982 Population Census: Main Demographic Features of the Population of Zimbabwe: An Advance Report Based on a Ten Percent Sample.* Central Statistical Office, Harare.

24. Chimombo, M. (1995). Malawi uses games to educate the young. *AIDS Analysis Africa (Southern Africa Edition)*, Vol. 6, No. 4, pp. 4.

25. Chin, J. and Lwanga, S. K. (1991). Estimation and projection of adult AIDS cases: a simple epidemiological model. *Bulletin of the World Health Organisation*, Vol. 69, Part 4, pp. 399-406.

26. Chin, J. and Mann, J. M. (1988). Global patterns and prevalence of AIDS and HIV infection. *AIDS*, Vol. 2, suppl. 1, pp. 247-252.

27. Chin, J., Remenyi, M. A., Morrison, F. and Bulatao, R. (1992). The global epidemiology of the HIV/AIDS pandemic and its projected demographic impact in Africa. *World Health Statistics Quarterly*, Vol. 45, pp. 220-227.

28. Clarke, G. M. and Cooke, D. (1992). *A Basic Course in Statistics*. Third Edition, Edward Arnold.

29. Coale, A. J. and Demeny, P. (1966). *Regional Model Life Tables and Stable Populations*. Princeton University Press, Princeton.

30. Cross, S. (1993). A socio-economic analysis of the long-run effects of AIDS in South Africa. In: Whiteside, A. and Cross, S., *Facing up to AIDS: The Socio-Economic Impact in Southern Africa*. St. Martin's Press, pp. 137-157.

31. CSS (1970). *Population Census 1970*. Report No. 02-05-01. Central Statistical Service, Pretoria.

32. CSS (1982). *Population Census 1980. Sample Tabulation*. Report No. 02-80-03. Central Statistical Service, Pretoria.

33. CSS (1986). *Standard Classification of Occupations*. Report No. 09-90-01, First Edition. Central Statistical Service, Pretoria.

34. CSS (1987). *South African Life Tables: 1984-1986*. Report No. 02-06-04. Central Statistical Service, Pretoria.

35. CSS (1991 a). *Deaths - 1991*. Report No. 03-09-01. Central Statistical Service, Pretoria.

36. CSS (1991 b). *Population Census 1991. Adjustment for Under-count*. Report No. 03-01-26. Central Statistical Service, Pretoria.

37. CSS (1991 c). *Population Census 1991. Age by Development Region, Statistical Region and District*. Report No. 03-01-03. Central Statistical Service, Pretoria.

38. CSS (1991 d). *Population Census 1991. Economic Characteristics of the Population*. Report No. 03-01-23. Central Statistical Service, Pretoria.

39. CSS (1992). *Recorded Deaths 1992*. Report No. 03-09-01. Central Statistical Service, Pretoria.

40. CSS (1993 a). *Bulletin of Statistics: Quarter Ending December 1993*. Vol. 27 (4). Central Statistical Service, Pretoria.

41. CSS (1993 b). *Recorded Deaths 1993*. Report No. 03-09-01. Central Statistical Service, Pretoria.

42. CSS (unpublished). *Recorded Deaths 1994*. Central Statistical Service, Pretoria.

43. CSS (1995). *October Household Survey 1994*. Statistical release P0317. Central Statistical Service, Pretoria.

44. CSS (1996 a). *Bulletin of Statistics: Quarter Ending December 1996*. Vol. 30 (4). Central Statistical Service, Pretoria.

45. CSS (1996 b). *October Household Survey 1995*. Statistical release P0317. Central Statistical Service, Pretoria.

46. CSS (1997). *Census '96: Preliminary Estimates of the Size of the Population of South Africa*. Central Statistical Service, Pretoria.

47. Department of Health (1991). First national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1990. *Epidemiological Comments*, Vol. 18, No. 2, pp. 23-45.

48. Department of Health (1992). Second national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1991. *Epidemiological Comments*, Vol. 19, No. 5, pp. 80-92.

49. Department of Health (1993). Third national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1992. *Epidemiological Comments*, Vol. 20, No. 3, pp. 35-47.

50. Department of Health (1994). Fourth national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1993. *Epidemiological Comments*, Vol. 21, No. 4, pp. 68-78.

51. Department of Health (1995 a). Fifth national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1994. *Epidemiological Comments*, Vol. 22, No. 5, pp. 90-100.

52. Department of Health (1995 b). AIDS in South Africa. Reported AIDS cases as on 30 November 1995. *Epidemiological Comments*, Vol. 22, No. 10, pp. 233.

53. Department of Health (1996). Sixth national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1995. *Epidemiological Comments*, Vol. 23, No. 1, pp. 3-17.

54. Department of Health (1996/1997). Seventh national HIV survey of women attending antenatal clinics of the public health services in the Republic of South Africa, October/November 1996. *Epidemiological Comments*, Vol. 23, No. 2, pp. 4-16.

55. Directorate of Statistics (undated). *1991 Population Census of the Former Bophuthatswana. Analytical Report: Nuptiality, Mortality and Fertility Patterns*. Vol. 2, Part 3. Mafikeng, North West Province.

56. Doehring, R. O. (1991). *The Socio-Economic Impact of the AIDS Epidemic.* MBA Dissertation, University of the Witwatersrand, Johannesburg.

57. Dorrington, R. E. (1989). African mortality rates - an initial estimate. *Transactions of the Actuarial Society of South Africa*, Vol. VII, Part I, pp. 169-202.

58. Dorrington, R. E., Martens, E. and Slawski, J. (1991). African mortality in South Africa: Assured lives, members of group schemes, and the population as a whole. *Transactions of the Actuarial Society of South Africa*, Vol. IX, Part I, pp. 268-288.

59. Doyle, P. R. and Millar, D. B. (1990). A general description of an actuarial model applicable to the HIV epidemic in South Africa. *Transactions of the Actuarial Society of South Africa*. Vol. VIII, Part II, pp. 561-593.

60. Doyle, P. R. (1991). *AIDS in South Africa: The Demographic and Economic Implications. Part I: The Impact of AIDS on the South African Population.* Paper No. 23. The Centre for Health Policy, Department of Community Health, Medical School, University of the Witwatersrand, Johannesburg.

61. Doyle, P. R. (1996). *Reaction to the Results of the Sixth National HIV Survey.* Metropolitan Life, unpublished.

62. Doyle, P. R., Mühr, T., Steinberg, M. and Broomberg, J. (1997). Towards optimisation of the HIV/AIDS modelling process in South Africa. *The Socio-Demographic Impact of AIDS in Africa Conference*, International Union for the Scientific Study of Population, University of Natal. Durban, 3-6 Feb. 1997.

63. Edelston, K. (1988). *AIDS: Countdown to Doomsday.* Media House Publications, Johannesburg.

64. FitzSimons, D. W. (1993). The global pandemic of AIDS. In: Cross, S. and Whiteside, A., *Facing up to AIDS: The Socio-Economic Impact in Southern Africa*. St. Martin's Press, pp. 13-33.

65. Fleming, A. (1993). Lessons from tropical Africa for addressing the HIV/AIDS epidemic in South Africa. In: Cross, S. and Whiteside, A., *Facing up to AIDS: The Socio-Economic Impact in Southern Africa*. St. Martin's Press, pp. 295-317.

66. Foster, H. W. and Keen, A. W. (1988). Black attitudes in suicide. In: Schlebusch, L. (1988). *Suicidal behaviour 2*. Proceedings of the Second Southern African Conference on Suicidology, 2-3 September, Durban, South Africa.

67. Foster, S. (1996). The implications of HIV/AIDS for South African Mines. *AIDS Analysis Africa (Southern Africa Edition)*, Vol. 7, No. 3, pp. 5.

68. Greenberg, A. E. *et al* (1995). Autopsy-proven causes of death in HIV-infected patients treated for tuberculosis in Abidjan, Côte d'Ivoire. *AIDS*, Vol. 9, pp. 1251-1254.

69. Gregson, S., Zaba, B., Garnet, G. and Anderson, R. (1997). Projecting the HIV/AIDS epidemic in Southern Africa. *The Socio-Demographic Impact of AIDS in Africa Conference*, International Union for the Scientific Study of Population, University of Natal. Durban, 3-6 Feb. 1997.

70. Groeneveld, H. (1990). *A stochastic micro-simulation model of the HIV epidemic.* Medical Research Council lecture, 23 March 1990.

71. Groeneveld, H. and Padayachee, G. N. (1992). A stochastic model for medium-term estimation of the prevalence of HIV infection in a South African heterosexual population. *South African Medical Journal*, Vol. 81, pp. 67-70.

72. Hambridge, M. (1990). AIDS/HIV and social dislocation in Natal. *AIDS Analysis Africa (Southern Africa Edition)*, Vol. 1, No. 3, pp. 1-2.

73. Heligman, L. and Pollard, J. H. (1980). The age pattern of mortality. *Journal*

of the Institute of Actuaries, Vol. 107, Part 1, pp. 49-80.

74. Hessol, N. A., Buchbinder, S. P., Colbert, D., Scheer, S., Underwood, R., Barnhart, L., O'Malley, P. M., Doll, L. S. and Lifson, A. R. (1992). *American Journal of Public Health*, Vol. 82 (4), pp. 561-564.

75. Hill, K. (1990). *PROJ3S – A Computer Program for Population Projections*. The World Bank, Washington, DC.

76. Ijsselmuiden, C. (1992). AIDS in Africa. *Epidemiological Comments*, Vol. 19(12), pp. 211-215.

77. Jeffery, A. (1997). Sixteen years of strife. *Fast Facts*, South African Institute of Race Relations, June.

78. Jentsch, U. (1996). *HIV Related Mortality in South African Adult Patients*. M.Med Thesis, University of the Witwatersrand, Johannesburg.

79. Kielkowski, D., Steinberg, M. and Barron, P. M. (1989). Life after death - mortality statistics and the public health. *South African Medical Journal*, Vol. 76, pp. 672-675.

80. Lachman, S. J. (1995). *Heterosexual HIV/AIDS as a Global Problem. A Guide for Medical Practitioners and Health Care Workers*. Fifth Edition, TPS Drug Information Centre (Pty) Ltd.

81. Leroy, V. *et al* (1995). Four years of natural history of HIV-1 infection in African women: a prospective cohort study in Kigali (Rwanda), 1988-1993. *Journal of the Acquired Immune Deficiency Syndrome*, Vol. 9, pp. 415-421.

82. Lörper, J. (1989). Projecting the spread of AIDS into the general population - application to life assurance. *Journal of the Institute of Actuaries*, Vol. 116, pp. 625-638.

83. Lucas, S. B. *et al* (1993). The mortality and pathology of HIV infection in a West African city. *AIDS*, Vol. 7, pp. 1569-1579.

84. Mandela, N. (1997). AIDS: Facing up to the global threat. *World Economic Forum Annual Meeting*. Davos, Switzerland, February 1997.

85. Mandevu, R. (1995). Young girls become "easy prey" in Botswana. *AIDS Analysis Africa (Southern Africa Edition)*, Vol. 6, No. 2, pp. 9-10.

86. Manton, K. G., Poss, S. S. and Wing, S. (1979). The black/white mortality crossover: Investigation from the perspectives of the components of ageing. *Gerontologist*, Vol. 19, pp. 291-299.

87. Manton, K. G. and Stallard, E. (1984). *Recent Trends in Mortality Analysis*. Centre for Demographic Studies, Duke University. Academic Press, Inc..

88. Mayekiso, T. V. (1995). Attitudes of black adolescents towards suicide. In: Schlebusch, L. (1995). *Suicidal Behaviour 3*. Proceedings of the Third Southern African Conference on Suicidology, 12-13 May, Durban, South Africa.

89. Miller, N. and Yeager, R. (1996). AIDS and the military. *AIDS Analysis Africa (Southern Africa Edition)*, Vol. 6, No. 4, pp. 5-6.

90. Mostert, W. P. and Hofmeyer, B. E. (in press). *Fertility, Mortality, Family Planning and Related Indicators in South Africa*. Human Sciences Research Council, Pretoria.

91. Nam, C. B., Weatherby, N. L. and Ockay, K. A. (1978). Causes of death which contribute to the mortality crossover effect. *Social Biology*, Vol. 25, pp. 306-314.

92. NCHS (National Centre for Health Statistics) (1975). United States life tables: 1969-1971. Department of Health, Education and Welfare Pub. No. (HRA) 75-1150. *U.S. Decennial Life Tables for 1969-1971*, Vol. 1, No. 1. Rockville, Maryland: Public Health Service.

93. Neill, A. (1977). *Life Contingencies*. The Institute of Actuaries and the Faculty of Actuaries in Scotland, Great Britain.

94. Nelson, A. M. *et al* (1993). A clinical and pathological comparison of the WHO and CDC case definitions for AIDS in Kinshasa, Zaire: is passive surveillance valid? *AIDS*, Vol. 7, 1241-1245.

95. Newell, C. (1994). *Methods and Models in Demography*. John Wiley and Sons.

96. Okware, S. I. (1987). Towards a national AIDS-control program in Uganda. *The Western Journal of Medicine*, Vol. 146, Part 6, pp. 726-729.

97. Over, M. (1990). *The Economic Impact of Fatal Adult Illness from AIDS and Other Causes in Sub-Saharan Africa: A Research Proposal*. Unpublished document, Research Department, World Bank, Washington.

98. Padayachee, G. N. and Schall, R. (1990). Short-term predictions of the prevalence of human immunodeficiency virus infection among the black population in South Africa. *South African Medical Journal*, Vol. 77, pp. 329-333.

99. Patel, L. (1995). *The Social and Economic Impact of HIV/AIDS on Individuals and Families*. Paper No. 40, Centre for Health Policy, Department of Community Health, University of the Witwatersrand, Johannesburg.

100. Perks, W. F. (1932). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, Vol. 63, pp. 12-40.

101. Phillips, A. N., Sabin, C. A., Elford, J., Bofill, M., Janossy, G., Lee and C. A. (1994). Use of CD4 lymphocyte count to predict long term survival free of AIDS after HIV infection. *British Medical Journal*, Vol. 309, pp. 303-313.

102. Preston, S., Coale, A. J., Trussell, J. and Weinstein, M. (1980). Estimating the completeness of reporting of adult deaths in populations that are approximately stable. *Population Index*, Vol. 46(2), pp. 179-202.

103. Sadie, J. L. (1973). *Projections of the South African Population 1970-2020*. Industrial Development Corporation.

104. Sadie, J. L. (1988). *A Reconstruction and Projection of Demographic Movements in the RSA and TBVC Countries*. Research Report No. 148. Bureau of Market Research, University of South Africa, Pretoria.

105. Sadie, J. L. (1993). *A Projection of the South African Population, 1991-2011*. Research Report No. 196. Bureau of Market Research, University of South Africa, Pretoria.

106. Schall, R. (1990). On the maximum size of the AIDS epidemic among the heterosexual black population in South Africa. *South African Medical Journal*, Vol. 78, pp. 507-510.

107. Schall, R. and Padayachee, G. N. (1990). Doomsday forecasts of the AIDS epidemic. *South African Medical Journal*, Vol. 78, pp. 503.

108. Schlemmer, L. and Levitz, C. (1997). A very serious underestimate? *Fast Facts*, August 1997, pp. 3-4. South African Institute of Race Relations.

109. Schneider, H. (1993). What is being done about STD control in South Africa? *AIDS Bulletin*, Vol. 2, No. 2, pp. 1-2.

110. Schoub, B. D., Smith, A. N., Johnson, S., Martin, D. J., Lyons, S. F., Padayachee, G. N. and Hurwitz, H. S. (1990). Considerations on the further expansion of the AIDS epidemic in South Africa - 1990. *South African Medical Journal*, Vol. 77, pp. 613-618.

111. Schoub, B. D. (1994). *AIDS and HIV in Perspective. A Guide to Understanding the Virus and its Consequences*. Cambridge University Press.

112.Sitas, F., Terblanche, M. and Madhoo, J. (1996). *Incidence and Geographical Distribution of Histologically Diagnosed Cancer in South Africa, 1990 and 1991*. National Cancer Registry, South African Institute for Medical Research, Johannesburg.

113.Sitas, F., Bezwoda, W. R., Levin, V., Ruff, P., Kew, M. C., Hale, M. J., Carrara, H., Beral, V., Fleming, G., Odes, R. and Weaving, A. (1997). Association between human immunodeficiency virus type 1 infection and cancer in the black population of Johannesburg and Soweto, South Africa. *British Journal of Cancer*, Vol. 75 (11), pp. 1704-1707.

114.Solomon, G. I. (1995). AIDS Memoire: Reportback of the Actuarial Society AIDS Committee. *Transactions of the Actuarial Society of South Africa*, Vol. X, Part III, pp. 250-270.

115.South African Institute of Race Relations (1997). *South Africa Survey 1996/97*. Johannesburg.

116.ter Meulen, J., Eberhardt, H. C., Luande, J., Myaga, H. N., Chang-Claude, J., Mtiro, H., Mhina, M., Kashaija, P., Ockert, S., Yu, X., Meinhardt, G., Gissman, L. and Pawlita, M. (1992). Human papilloma virus (HPV) infection, HIV infection and cervical cancer in Tanzania, East Africa. *International Journal of Cancer*, Vol. 54, pp. 26-36.

117.*The Star* (1997). HIV alarm as 2,5-million are found to be infected. 25 April 1997.

118.Udjo, E. O. (1995). Childhood and adult mortality patterns in Botswana. In: *1991 Population and Housing Census Dissemination Seminar*, 1-4 May 1995, Gaborone. Central Statistics Office, Gaborone.

119.Udjo, E. O. (1997). *Fertility and Mortality Trends in South Africa: The Evidence from the 1995 October Household Survey, and Implications on Population Projections*. Unpublished monograph.

120.UN (1982). *Unabridged Model Life Tables Corresponding to the New United Nations Model Life Tables for Developing Countries*. Department of International Economic and Social Affairs. United Nations, New York.

121.UN (1993). *Manual X. Indirect Techniques for Demographic Estimation*. Department of International and Social Affairs. Population Studies Report No. 81. United Nations, New York.

122.UN (1994). *AIDS and the Demography of Africa*. United Nations, New York.

123.UNAIDS (1996). *HIV/AIDS: The Global Epidemic December 1996*. The Joint

United Nations Programme on HIV/AIDS. Website http://www.unaids.org/highband/document/epidemio/situat96.html.

124.UNAIDS (1997). *Global HIV/AIDS Epidemic December 1997.* The Joint United Nations Programme on HIV/AIDS. Website http://www.unaids.org/highband/document/epidemio/report97.html.

125.USGAO (1989). *AIDS Forecasting: Undercount of Cases and Lack of Key Data Weaken Existing Estimates.* US General Accounting Office, Washington DC.

126.van der Linde, I. (1996). Condoms in SA. *South African Medical Journal,* Vol. 86, No. 11, pp. 1468.

127.van der Merwe, S., Yach, D. and Metcalf, C. A. (1991). Peering into the black hole - the quality of black mortality data in Port Elizabeth and the rest of South Africa. *South African Medical Journal,* Vol. 79, pp. 419-422.

128.Whiteside, A. (1993). *AIDS: Socio-Economic Causes and Consequences.* Occasional Paper No. 29, Economic Research Unit, University of Natal, Durban.

129.Whiteside, A. (1995). And now for some good news. *AIDS Analysis Africa (Southern Africa Edition),* Vol. 6, No. 3, pp. 5-6.

130.WHO (1975). *Ninth Revision of the International Statistical Classification of Diseases,* Injuries and Causes of Death. World Health Organisation, Geneva.

131.WHO (1988). *World Health Statistics Annual.* World Health Organisation, Geneva.

132.WHO (1989). *Projection of HIV/AIDS in Sub-Saharan Africa Using an Epidemiologically Based Modelling Approach.* World Health Organisation, Geneva.

133.WHO (1992). *Tenth Revision of the International Statistical Classification of Diseases and Related Health Problems.* World Healtn Organisation, Geneva.

134.WHO (1994). *AIDS: Images of the Epidemic.* World Health Organisation, Geneva.

135.Zaba, B. (1981). *Use of the Relational Gompertz Model in Analysing Fertility Data Collected in Retrospective Surveys.* Centre for Population Studies Working Paper 8. London School of Hygiene and Tropical Medicine, University of London.

BIBLIOGRAPHY

1. Barnett, T. and Blaikie, P. (1992). *AIDS in Africa: Its Present and Future Impact*. Belhaven Press, London.

2. Brookmeyer, R. and Damiano, A. (1989). Statistical methods for short-term projections of AIDS incidence. *Statistics in Medicine*, Vol. 8, pp. 23-34.

3. Cross, M. and Moscardini, A. O. (1985). *Learning the Art of Mathematical Modelling*. Ellis Horwood Limited, Chichester.

4. Cross, S. and Whiteside, A. (1993). *Facing up to AIDS: The Socio-Economic Impact in Southern Africa*. St. Martin's Press.

5. Fowkes, N. D. and Mahony, J. J. (1994). *An Introduction to Mathematical Modelling*. John Wiley and Sons.

6. Golub, A., Gorr, W. L. and Gould, P. R. (1993). Spatial diffusion of the HIV/AIDS epidemic: modelling implications and case study of AIDS incidence in Ohio. *Geographical Analysis*, Vol. 25(2), pp. 85-100.

7. Isham, V. (1988). Mathematical modelling of the transmission dynamics of HIV infection and AIDS: a review. *Journal of the Royal Statistical Society A*, Vol. 151, Part 1, pp. 5-30.

8. Kapur, J. N. (1988). *Mathematical Modelling*. John Wiley and Sons.

9. Lachman, S. J. (1995). *Heterosexual HIV/AIDS as a Global Problem. A Guide for Medical Practitioners and Health Care Workers*. Fifth Edition, TPS Drug Information Centre (Pty) Ltd.

10. Rusenberg, S., Cooke, K. and Hsien, Y. H. (1995). A model for HIV in Asia. *Mathematical Biosciences*, Vol. 128, pp. 185-210.

11. Schoub, B. D. (1994). *AIDS and HIV in Perspective. A Guide to Understanding the Virus and its Consequences.* Cambridge University Press.

12. Tan, W. Y. and Hsu, H. (1989). Some stochastic models of AIDS spread. *Statistics in Medicine*, Vol. 8, pp. 121-136.

13. Thomas, R. (1996). Modelling space-time HIV/AIDS dynamics: applications to disease control. *Social Science and Medicine*, Vol. 43(3), pp. 353-366.

14. WHO (1994). *AIDS: Images of the Epidemic.* World Health Organisation, Geneva.

15. Wiley, J. A., Herschkorn, S. J. and Padian, N. S. (1989). Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penile-vaginal intercourse. *Statistics in Medicine*, Vol. 8, pp. 93-102.

16. Zeger, S. L., See, L. and Diggle, P. J. (1989). Statistical methods for monitoring the AIDS epidemic. *Statistics in Medicine*, Vol. 8, pp. 3-21.

Author Mitchell J A
Name of thesis The Estimation Of Aids-Related Deaths Amongst The Black Population Of South Africa And The Impact On The Assa Aids Model Mitchell J A 1998

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.